

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:54:17 ; Search time 304 Seconds

(without alignments)  
1340.366 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMNLVNRNLEFQVLRVRE.....SKSATNIGRQGFASPMK 632

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US111\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3243	100.0	632	21	US-09-791-537-1796
2	3243	100.0	632	22	US-09-881-736-2
3	3237	99.8	632	20	US-09-641-377-585
4	3233	99.7	632	18	US-09-488-725A-3192
5	3230	99.6	632	21	US-09-791-537-72082
6	3230	99.6	632	22	US-09-833-790-413

7	2919	90.0	570	22	US-09-833-790-427	Sequence 427, App
8	2734	84.3	628	21	US-09-791-537-36304	Sequence 36304, A
9	2734	84.3	628	22	US-09-881-736-4	Sequence 4, App1
10	1182	36.4	255	20	US-09-628-468A-14587	Sequence 14587, A
11	1182	36.4	255	27	US-60-389-987-1045	Sequence 1045, App
12	1182	36.4	255	27	US-60-412-418-1045	Sequence 1045, App
13	1103.5	34.0	243	21	US-09-758-472-6751	Sequence 6751, App
14	1103.5	34.0	243	26	US-10-235-926-6751	Sequence 6751, App
15	971.5	30.0	256	1	PCT-US01-08656-9072	Sequence 9072, App
16	940.5	29.0	201	27	US-60-182-569-1026	Sequence 1026, App
17	889.5	27.4	191	27	US-60-182-569-1025	Sequence 1025, App
18	864	26.6	625	20	US-09-614-150-30960	Sequence 30960, A
19	864	26.6	625	21	US-09-791-537-49124	Sequence 49124, A
20	864	26.6	625	27	US-60-191-637-30526	Sequence 30526, A
21	749.5	23.1	165	27	US-60-186-662-967	Sequence 967, App
22	662.5	20.4	148	27	US-60-186-662-968	Sequence 968, App
23	646	19.9	187	27	US-60-207-215-294	Sequence 294, App
24	632	19.5	142	27	US-60-186-655-696	Sequence 696, App
25	631	19.5	872	1	PCT-US01-08656-9071	Sequence 9071, App
26	631	19.5	872	27	US-60-311-261-2570	Sequence 2570, App
27	625	19.3	4318	1	PCT-US01-08631-45146	Sequence 45146, A
28	618	19.1	665	1	PCT-US01-08631-40589	Sequence 40589, A
29	618	19.1	1086	1	PCT-US01-08631-50263	Sequence 50263, A
30	618	19.1	1139	1	PCT-US01-08631-44317	Sequence 44317, A
31	581	17.9	119	18	US-09-488-725A-6764	Sequence 6764, App
32	549	16.9	135	27	US-60-162-245-3866	Sequence 3866, App
33	549	16.9	135	27	US-60-162-245-3866	Sequence 3866, App
34	547.5	16.9	135	27	US-60-162-245-3866	Sequence 3866, App
35	541.5	16.7	881	21	US-09-791-537-76252	Sequence 76252, A
36	521	16.1	114	27	US-60-162-245-3547	Sequence 3547, App
37	477	14.7	383	21	US-09-791-537-35867	Sequence 35867, A
38	477	14.7	383	21	US-09-791-537-77334	Sequence 77334, A
39	477	14.7	384	20	US-09-614-150-4533	Sequence 4533, App
40	477	14.7	384	21	US-09-791-537-65301	Sequence 65301, App
41	477	14.7	384	27	US-60-167-217-4632	Sequence 4632, App
42	477	14.7	384	27	US-60-191-637-4551	Sequence 4551, App
43	477	14.7	385	21	US-09-791-537-150251	Sequence 150251, App
44	460	14.2	367	27	US-60-146-315-901	Sequence 901, App
45	334	10.3	2022	27	US-60-389-987-558	Sequence 558, App

## ALIGNMENTS

RESULT 1  
US-09-791-537-1796  
; Sequence 1796, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791, 537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1796  
; LENGTH: 632  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-1796

Query Match	100.0%	Score 3243;	DB 21;	Length 632;
Best Local Similarity	100.0%	Pred. No. 2.8e+263;		
Matches 632;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps 0;	
QY	1	MDTMNLVNRNLEFQVLRVREISSENEVOFIOADDFEDFRKKWORTHELKRYDLMK	60	
DB	1	MDTMNLVNRNLEFQVLRVREISSENEVOFIOADDFEDFRKKWORTHELKRYDLMK	60	
QY	61	AETERSALDVKIKHARNQVDVEIKRRQREACCEKLERQIQIIRLMQDTSIGSIQISPE	120	

```
Db 61 AETERSALDVKLKHARNOVVEIKRORADECEKLEKROJOLIREMIMCTSGSIQJSEE 120
Oy 121 OKSALAFNLNGOPSSSAGNKRRLSTIDESSIISDLSFDEKTDSDSLMDSSLVKTFKLR 180
Db 121 OKSALAFNLNGOPSSSAGNKRRLSTIDESSIISDLSFDEKTDSDSLMDSSLVKTFKLR 180
Oy 181 EKRRSTSRQFVDPGPGVKKTRISGSAVDGNSIYAKTVTVVNDGPIEAVSTIETVP 240
Db 181 EKRRSTSRQFVDPGPGVKKTRISGSAVDGNSIYAKTVTVVNDGPIEAVSTIETVP 240
Oy 241 YWTRSRKKTGTLOPWNDSSTLNSRQLEPRRETDSVGTPOSGMRLHDFVSKTYIKPESC 300
Db 241 YWTRSRKKTGTLOPWNDSSTLNSRQLEPRRETDSVGTPOSGMRLHDFVSKTYIKPESC 300
Oy 301 VPGGKRIFGKLSLKDCDCHVSHPECRDRCPLPCPTLTIGTPVKIGEGMLADPVSOTSP 360
Db 301 VPGGKRIFGKLSLKDCDCHVSHPECRDRCPLPCPTLTIGTPVKIGEGMLADPVSOTSP 360
Oy 361 MIRSIVHVCNEIEORGLTETGLYRISGCDRTYKELKEKFLRVKTVPLLSKVDIHAICS 420
Db 361 MIRSIVHVCNEIEORGLTETGLYRISGCDRTYKELKEKFLRVKTVPLLSKVDIHAICS 420
Oy 421 LKDFLNLKEPLTFPLNAPFMAEITTEDNSIAAMYAVGELPOANDDTLAFMLIHL 480
Db 421 LKDFLNLKEPLTFPLNAPFMAEITTEDNSIAAMYAVGELPOANDDTLAFMLIHL 480
Oy 481 ORVAOSPHTKMDVANLAKVGPITVAHAVNPDPVTMSODIKROPKVERLLSLPLEYWS 540
Db 481 ORVAOSPHTKMDVANLAKVGPITVAHAVNPDPVTMSODIKROPKVERLLSLPLEYWS 540
Oy 541 QFMWVEQENIDPLHVIENSNAFSTPQTPDIKVSILGVTTPHEHQLTPSSSSLSQVRVS 600
Db 541 QFMWVEQENIDPLHVIENSNAFSTPQTPDIKVSILGVTTPHEHQLTPSSSSLSQVRVS 600
Oy 601 TLTKNTPREGSKSKSATNLGRQGNFASPMUK 632
Db 601 TLTKNTPREGSKSKSATNLGRQGNFASPMUK 632
```

## RESULT 2

```
US-09-881-736-2
; Sequence 2, Application US/09881736
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: Cyt-4 polypeptides, DNA molecules encoding them and their use in
; FILE REFERENCE: 0652.2260001/EKS/AES
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-881-736-2
```

Query Match 100.0%; Score 3243; DB 22; Length 632;  
Best Local Similarity 100.0%; Pred. No. 2,8e-263;  
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 MDTMMNLNRLNLFQOLYRVEILSEGNEVQFIQAKDFEDERKKMORTDHELGRYKDLAM 60
Db 1 MDTMMNLNRLNLFQOLYRVEILSEGNEVQFIQAKDFEDERKKMORTDHELGRYKDLAM 60
Oy 61 AETERSALDVKLKHARNOVVEIKRORADECEKLEKROJOLIREMIMCTSGSIQJSEE 120
Db 61 AETERSALDVKLKHARNOVVEIKRORADECEKLEKROJOLIREMIMCTSGSIQJSEE 120
Oy 121 OKSALAFNLNGOPSSSAGNKRRLSTIDESSIISDLSFDEKTDSDSLMDSSLVKTFKLR 180
Db 121 OKSALAFNLNGOPSSSAGNKRRLSTIDESSIISDLSFDEKTDSDSLMDSSLVKTFKLR 180
Oy 181 EKRRSTSRQFVDPGPGVKKTRISGSAVDGNSIYAKTVTVVNDGPIEAVSTIETVP 240
Db 181 EKRRSTSRQFVDPGPGVKKTRISGSAVDGNSIYAKTVTVVNDGPIEAVSTIETVP 240
Oy 241 YWTRSRKKTGTLOPWNDSSTLNSRQLEPRRETDSVGTPOSGMRLHDFVSKTYIKPESC 300
Db 241 YWTRSRKKTGTLOPWNDSSTLNSRQLEPRRETDSVGTPOSGMRLHDFVSKTYIKPESC 300
Oy 301 VPGGKRIFGKLSLKDCDCHVSHPECRDRCPLPCPTLTIGTPVKIGEGMLADPVSOTSP 360
Db 301 VPGGKRIFGKLSLKDCDCHVSHPECRDRCPLPCPTLTIGTPVKIGEGMLADPVSOTSP 360
Oy 361 MIRSIVHVCNEIEORGLTETGLYRISGCDRTYKELKEKFLRVKTVPLLSKVDIHAICS 420
Db 361 MIRSIVHVCNEIEORGLTETGLYRISGCDRTYKELKEKFLRVKTVPLLSKVDIHAICS 420
Oy 421 LKDFLNLKEPLTFPLNAPFMAEITTEDNSIAAMYAVGELPOANDDTLAFMLIHL 480
Db 421 LKDFLNLKEPLTFPLNAPFMAEITTEDNSIAAMYAVGELPOANDDTLAFMLIHL 480
Oy 481 ORVAOSPHTKMDVANLAKVGPITVAHAVNPDPVTMSODIKROPKVERLLSLPLEYWS 540
Db 481 ORVAOSPHTKMDVANLAKVGPITVAHAVNPDPVTMSODIKROPKVERLLSLPLEYWS 540
Oy 541 QFMWVEQENIDPLHVIENSNAFSTPQTPDIKVSILGVTTPHEHQLTPSSSSLSQVRVS 600
Db 541 QFMWVEQENIDPLHVIENSNAFSTPQTPDIKVSILGVTTPHEHQLTPSSSSLSQVRVS 600
Oy 601 TLTKNTPREGSKSKSATNLGRQGNFASPMUK 632
Db 601 TLTKNTPREGSKSKSATNLGRQGNFASPMUK 632
```

## RESULT 3

```
US-09-641-377-585
; Sequence 585, Application US/09641377
; GENERAL INFORMATION:
; APPLICANT: WIEGMANN, STEFAN
; APPLICANT: GASENHUBER, JOHANN
; APPLICANT: TAMPE, JENS
; TITLE OF INVENTION: HUMAN DNA SEQUENCES
; FILE REFERENCE: 067100/0106
; CURRENT APPLICATION NUMBER: US/09/641,377
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,499
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/156,503
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 1793
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 585
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-641-377-585
```

Query Match 99.8%; Score 3237; DB 20; Length 632;  
Best Local Similarity 99.8%; Pred. No. 8.9e-263;  
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MOTMLANVNLFEOLVRRVEILLSEGEVQFIOLAKDFEDFRKKWORTHELKGYKDLMM 60
DB 1 MOTMLANVNLFEOLVRRVEILLSEGEVQFIOLAKDFEDFRKKWORTHELKGYKDLMM 60
QY 61 AETESALDVKLKHARNVDVEIKRQRAEADCEKLEKQIOLIRMLMCDISGSIQJLSEE 120
DB 61 AETESALDVKLKHARNVDVEIKRQRAEADCEKLEKQIOLIRMLMCDISGSIQJLSEE 120
QY 121 OKSALAFINRGQSSSSNAGNKRSLSTIDSGSLSDISFDKDESLDMSLSLVKTPFKLKR 180
DB 121 OKSALAFINRGQSSSSNAGNKRSLSTIDSGSLSDISFDKDESLDMSLSLVKTPFKLKR 180
QY 181 EKRSTSHQFDGPPGPKKTRISGSAVDGNESTVAKTQTVVPNDGPIEAVSTIETVP 240
DB 181 EKRSTSHQFDGPPGPKKTRISGSAVDGNESTVAKTQTVVPNDGPIEAVSTIETVP 240
QY 241 YWTSRRKTGTLOPWNSTLNSROLERPTETDSVGTPOSNGMRLHDFVSKTVIKPESC 300
DB 241 YWTSRRKTGTLOPWNSTLNSROLERPTETDSVGTPOSNGMRLHDFVSKTVIKPESC 300
QY 301 VPCGRIKFGKLSKRCRCRVVSHPECRCRPLPCIPILIGTPVKIGGMLADFYQSOTSP 360
DB 301 VPCGRIKFGKLSKRCRCRVVSHPECRCRPLPCIPILIGTPVKIGGMLADFYQSOTSP 360
QY 361 MIPSIVHCVNEIEBQRLTEGLYRISGCDRTVKELEKFLKVKTVPLLSKYDDIHAICS 420
DB 361 MIPSIVHCVNEIEBQRLTEGLYRISGCDRTVKELEKFLKVKTVPLLSKYDDIHAICS 420
QY 421 LKQFLNMLKEPLTLFRLNRAFMEEAETTEDNSIAAMYQAVGELPOANRDTLAFMLIHL 480
DB 421 LKQFLNMLKEPLTLFRLNRAFMEEAETTEDNSIAAMYQAVGELPOANRDTLAFMLIHL 480
QY 481 ORVASPHTKMDVANLAKVFGPTIVAHAVPNPDPTMSODIKRQKVERRLSLPLEYWS 540
DB 481 ORVASPHTKMDVANLAKVFGPTIVAHAVPNPDPTMSODIKRQKVERRLSLPLEYWS 540
QY 541 QPMWEOENIDPLHVIENSNAFSTPQTDIKVSLGPTTPEHQLKTPSSSSLSQRYRS 600
DB 541 QPMWEOENIDPLHVIENSNAFSTPQTDIKVSLGPTTPEHQLKTPSSSSLSQRYRS 600
QY 601 TLTKNTPRFGSKSKSATNMGROGNFASPMUK 632
DB 601 TLTKNTPRFGSKSKSATNMGROGNFASPMUK 632

RESULT 4
US-09-488-725A-3192
: Sequence 3192, Application US/09488725A
: GENERAL INFORMATION:
: APPLICANT: Hyseq Inc
: TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
: FILE REFERENCE: 784FLPCT
: CURRENT APPLICATION NUMBER: US/09/488,725A
: CURRENT FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US/09/488,725
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: US09/598,042
: PRIOR FILING DATE: 2000-06-20
: PRIOR APPLICATION NUMBER: US09/620,312
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: US09/653,450
: PRIOR FILING DATE: 2000-08-31
: PRIOR APPLICATION NUMBER: US09/662,191
: PRIOR FILING DATE: 2000-09-14
: PRIOR APPLICATION NUMBER: US09/693,036
: PRIOR FILING DATE: 2000-10-19
: PRIOR APPLICATION NUMBER: US09/727,344
: PRIOR FILING DATE: 2000-11-29
: NUMBER OF SEQ ID NOS: 7144
: SOFTWARE: PC_Fl_genes_b Versions 1.0
: SEQ ID NO 3192

```

```

: LENGTH: 633
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-488-725A-3192

Query Match          99.7%; Score 3233; DB 18; Length 632;
Best Local Similarity 99.7%; Pred. No. 1,9e+262;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MOTMLANVNLFEOLVRRVEILLSEGEVQFIOLAKDFEDFRKKWORTHELKGYKDLMM 60
DB 1 MOTMLANVNLFEOLVRRVEILLSEGEVQFIOLAKDFEDFRKKWORTHELKGYKDLMM 60
QY 61 AETESALDVKLKHARNVDVEIKRQRAEADCEKLEKQIOLIRMLMCDISGSIQJLSEE 120
DB 61 AETESALDVKLKHARNVDVEIKRQRAEADCEKLEKQIOLIRMLMCDISGSIQJLSEE 120
QY 121 OKSALAFINRGQSSSSNAGNKRSLSTIDSGSLSDISFDKDESLDMSLSLVKTPFKLKR 180
DB 121 OKSALAFINRGQSSSSNAGNKRSLSTIDSGSLSDISFDKDESLDMSLSLVKTPFKLKR 180
QY 181 EKRSTSHQFDGPPGPKKTRISGSAVDGNESTVAKTQTVVPNDGPIEAVSTIETVP 240
DB 181 EKRSTSHQFDGPPGPKKTRISGSAVDGNESTVAKTQTVVPNDGPIEAVSTIETVP 240
QY 241 YWTSRRKTGTLOPWNSTLNSROLERPTETDSVGTPOSNGMRLHDFVSKTVIKPESC 300
DB 241 YWTSRRKTGTLOPWNSTLNSROLERPTETDSVGTPOSNGMRLHDFVSKTVIKPESC 300
QY 301 VPCGRIKFGKLSKRCRCRVVSHPECRCRPLPCIPILIGTPVKIGGMLADFYQSOTSP 360
DB 301 VPCGRIKFGKLSKRCRCRVVSHPECRCRPLPCIPILIGTPVKIGGMLADFYQSOTSP 360
QY 361 MIPSIVHCVNEIEBQRLTEGLYRISGCDRTVKELEKFLKVKTVPLLSKYDDIHAICS 420
DB 361 MIPSIVHCVNEIEBQRLTEGLYRISGCDRTVKELEKFLKVKTVPLLSKYDDIHAICS 420
QY 421 LKQFLNMLKEPLTLFRLNRAFMEEAETTEDNSIAAMYQAVGELPOANRDTLAFMLIHL 480
DB 421 LKQFLNMLKEPLTLFRLNRAFMEEAETTEDNSIAAMYQAVGELPOANRDTLAFMLIHL 480
QY 481 ORVASPHTKMDVANLAKVFGPTIVAHAVPNPDPTMSODIKRQKVERRLSLPLEYWS 540
DB 481 ORVASPHTKMDVANLAKVFGPTIVAHAVPNPDPTMSODIKRQKVERRLSLPLEYWS 540
QY 541 QPMWEOENIDPLHVIENSNAFSTPQTDIKVSLGPTTPEHQLKTPSSSSLSQRYRS 600
DB 541 QPMWEOENIDPLHVIENSNAFSTPQTDIKVSLGPTTPEHQLKTPSSSSLSQRYRS 600
QY 601 TLTKNTPRFGSKSKSATNMGROGNFASPMUK 632
DB 601 TLTKNTPRFGSKSKSATNMGROGNFASPMUK 632

RESULT 5
US-09-791-537-72082
: Sequence 72082, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 72082
: LENGTH: 632
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-791-537-72082

```

Query Match 99.6%; Score 3230; DB 21; Length 632;  
Best Local Similarity 99.7%; Pred. No. 3.5e-262;  
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MDTMLNVRNLFQOLYARVEILSEGENVOPIQLAKDEFEDRKKWQRTDHLGKYLMMK 60
DB 1 MDTMLNVRNLFQOLYARVEILSEGENVOPIQLAKDEFEDRKKWQRTDHLGKYLMMK 60
QY 61 AETERSALDYKLNHARQOVVEIKRORADEKLEKROQLIREMLMCTSSIOJSEE 120
DB 61 AETERSALDYKLNHARQOVVEIKRORADEKLEKROQLIREMLMCTSSIOJSEE 120
QY 121 OKSALAFNLNGOPSSNAGNKRRLSTIDESSILSDISFDDTDSLDMDSSLVYTFKLKR 180
DB 121 OKSALAFNLNGOPSSNAGNKRRLSTIDESSILSDISFDDTDSLDMDSSLVYTFKLKR 180
QY 181 EKRRSTSRQFVDGPPGCVKKTSTISGAVDGNESYAKTIVYVNDGPIEAVSTIETVP 240
DB 181 EKRRSTSRQFVDGPPGCVKKTSTISGAVDGNESYAKTIVYVNDGPIEAVSTIETVP 240
QY 241 YWTRSRKKTGTLPWNDSSTLNSROLEPRTETDSVGTPOSNGGRLHDFVSKTYIKPESC 300
DB 241 YWTRSRKKTGTLPWNDSSTLNSROLEPRTETDSVGTPOSNGGRLHDFVSKTYIKPESC 300
QY 301 VPCGKRIFGKLSLKCDRCRVVSHPECRDRCPLCIPITLLIGTPVKIGEGMLADVFVSQTS 360
DB 301 VPCGKRIFGKLSLKCDRCRVVSHPECRDRCPLCIPITLLIGTPVKIGEGMLADVFVSQTS 360
QY 361 MIPSIYVHCNVEIQRRLTETGLYRISGCRITYKELKRLRYKTYVPLSKVDIHAICS 420
DB 361 MIPSIYVHCNVEIQRRLTETGLYRISGCRITYKELKRLRYKTYVPLSKVDIHAICS 420
QY 421 LKDFLNKLEPLTFPLNRAFMFAEITTEDNSIAMYOAVGELPOANDTLAFIMHIL 480
DB 421 LKDFLNKLEPLTFPLNRAFMFAEITTEDNSIAMYOAVGELPOANDTLAFIMHIL 480
QY 481 ORVAQSPHTKMDVANLAKVFGPTIVAHAVNPDPVTMSODIKRQPKVERLLSLPLEYWS 540
DB 481 ORVAQSPHTKMDVANLAKVFGPTIVAHAVNPDPVTMSODIKRQPKVERLLSLPLEYWS 540
QY 541 QFMWVEQENTDPLHVIENSNAFSTPOTPDIKVSLGVTTPPEHOLKTPSSSSLSQVR 600
DB 541 QFMWVEQENTDPLHVIENSNAFSTPOTPDIKVSLGVTTPPEHOLKTPSSSSLSQVR 600
QY 601 TLTKNTPRFGSKSKSATNLGROGNFASPMK 632
DB 601 TLTKNTPRFGSKSKSATNLGROGNFASPMK 632
```

RESULT 6  
US-09-833-790-413  
; Sequence 413, Application US/09833790  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Secrist, Heather  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Indirias, Carol Y.  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.512  
; CURRENT APPLICATION NUMBER: US/09/833,790  
; CURRENT FILING DATE: 2001-04-11  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 413  
; LENGTH: 632  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-790-413

Query Match 99.6%; Score 3230; DB 22; Length 632;  
Best Local Similarity 99.7%; Pred. No. 3.5e-262;  
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MDTMLNVRNLFQOLYARVEILSEGENVOPIQLAKDEFEDRKKWQRTDHLGKYLMMK 60
DB 1 MDTMLNVRNLFQOLYARVEILSEGENVOPIQLAKDEFEDRKKWQRTDHLGKYLMMK 60
QY 61 AETERSALDYKLNHARQOVVEIKRORADEKLEKROQLIREMLMCTSSIOJSEE 120
DB 61 AETERSALDYKLNHARQOVVEIKRORADEKLEKROQLIREMLMCTSSIOJSEE 120
QY 121 OKSALAFNLNGOPSSNAGNKRRLSTIDESSILSDISFDDTDSLDMDSSLVYTFKLKR 180
DB 121 OKSALAFNLNGOPSSNAGNKRRLSTIDESSILSDISFDDTDSLDMDSSLVYTFKLKR 180
QY 181 EKRRSTSRQFVDGPPGCVKKTSTISGAVDGNESYAKTIVYVNDGPIEAVSTIETVP 240
DB 181 EKRRSTSRQFVDGPPGCVKKTSTISGAVDGNESYAKTIVYVNDGPIEAVSTIETVP 240
QY 241 YWTRSRKKTGTLPWNDSSTLNSROLEPRTETDSVGTPOSNGGRLHDFVSKTYIKPESC 300
DB 241 YWTRSRKKTGTLPWNDSSTLNSROLEPRTETDSVGTPOSNGGRLHDFVSKTYIKPESC 300
QY 301 VPCGKRIFGKLSLKCDRCRVVSHPECRDRCPLCIPITLLIGTPVKIGEGMLADVFVSQTS 360
DB 301 VPCGKRIFGKLSLKCDRCRVVSHPECRDRCPLCIPITLLIGTPVKIGEGMLADVFVSQTS 360
QY 361 MIPSIYVHCNVEIQRRLTETGLYRISGCRITYKELKRLRYKTYVPLSKVDIHAICS 420
DB 361 MIPSIYVHCNVEIQRRLTETGLYRISGCRITYKELKRLRYKTYVPLSKVDIHAICS 420
QY 421 LKDFLNKLEPLTFPLNRAFMFAEITTEDNSIAMYOAVGELPOANDTLAFIMHIL 480
DB 421 LKDFLNKLEPLTFPLNRAFMFAEITTEDNSIAMYOAVGELPOANDTLAFIMHIL 480
QY 481 ORVAQSPHTKMDVANLAKVFGPTIVAHAVNPDPVTMSODIKRQPKVERLLSLPLEYWS 540
DB 481 ORVAQSPHTKMDVANLAKVFGPTIVAHAVNPDPVTMSODIKRQPKVERLLSLPLEYWS 540
QY 541 QFMWVEQENTDPLHVIENSNAFSTPOTPDIKVSLGVTTPPEHOLKTPSSSSLSQVR 600
DB 541 QFMWVEQENTDPLHVIENSNAFSTPOTPDIKVSLGVTTPPEHOLKTPSSSSLSQVR 600
QY 601 TLTKNTPRFGSKSKSATNLGROGNFASPMK 632
DB 601 TLTKNTPRFGSKSKSATNLGROGNFASPMK 632
```

RESULT 7  
US-09-833-790-427  
; Sequence 427, Application US/09833790  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Secrist, Heather  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Indirias, Carol Y.  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.512  
; CURRENT APPLICATION NUMBER: US/09/833,790  
; CURRENT FILING DATE: 2001-04-11  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 427  
; LENGTH: 570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-790-427

Query Match 90.0%; Score 2919; DB 22; Length 570;

Best Local Similarity	99.8%	Pred. No.	4.5e-236;
Matches	569;	Conservative	0;
		Mismatches	1;
		Indels	0;
		Gaps	0;

QY	63	TERALDYKLTAKHARQOVVEIKRRORAEADCKLEJOILJEMIMCPTSGS1QJLSEBOK	122
Dd	1	TERALDYKLTAKHARQOVVEIKRRORAEADCKLEJOILJEMIMCPTSGS1QJLSEBOK	60
QY	123	SALAFNLNGOPSSSSNAGNKRJLSTIDESGSILSDISFDKTDSELDWDSLVKTEFKLKRREK	182
Dd	61	SALAFNLNGOPSSSSNAGNKRJLSTIDESGSILSDISFDKTDSELDWDSLVKTEFKLKRREK	120
QY	183	RRSTRQOPVDPGPGVKKTRTSGSAVDQGNESIVAKTYTVTPNDGGRJEAIVSTIETVPYV	242
Dd	121	RRSTRQOPVDPGPGVKKTRTSGSAVDQGNESIVAKTYTVTPNDGGRJEAIVSTIETVPYV	180
QY	243	TRSRKKTGTLPMWNSDSTLNSROLBPRTTDSVGPQOSGMRJHDFVSKTYVKKPESCVP	302
Dd	181	TRSRKKTGTLPMWNSDSTLNSROLBPRTTDSVGPQOSGMRJHDFVSKTYVKKPESCVP	240
QY	303	CGKRIRKFGKLSLKCDKCDRVVSHPECRDRCPJLPCITLLGTVPKIGEGMADFEVSQSPMI	362
Dd	241	CGKRIRKFGKLSLKCDKCDRVVSHPECRDRCPJLPCITLLGTVPKIGEGMADFEVSQSPMI	300
QY	363	PSIYVHCYNELEQKGLTETGLYRISGCRHYAKELKEKFLRYKTVPLSKYVDIHAIICSLI	422
Dd	301	PSIYVHCYNELEQKGLTETGLYRISGCRHYAKELKEKFLRYKTVPLSKYVDIHAIICSLI	360
QY	423	KDPLRNLKEPLTTPFLNRAFMFAAITBEDNSIAMAQVAGELPQANRPTLAFLMHLQR	482
Dd	361	KDPLRNLKEPLTTPFLNRAFMFAAITBEDNSIAMAQVAGELPQANRPTLAFLMHLQR	420
QY	483	VAQSPHTKMDYANLAKVGPITYVAAYVNPDPVTASODIKRQPKYVERLSTLPLEYWSOF	542
Dd	421	VAQSPHTKMDYANLAKVGPITYVAAYVNPDPVTMHLQDKRQPKYVERLSTLPLEYWSOF	480
QY	543	MAVEQENTDPLAHYIENSNAFSTPQTPDIKVSLLGVPYTPREHOLKTPSSSSLSQRYRSTL	602
Dd	481	MAVEQENTDPLAHYIENSNAFSTPQTPDIKVSLLGVPYTPREHOLKTPSSSSLSQRYRSTL	540
QY	603	TKNTPRGSKSKSATNLGRQGNFASPMUX	632
Dd	541	TKNTPRGSKSKSATNLGRQGNFASPMUX	570

```

RESULT 8
US-09-791-537-36304
? Sequence 36304, Application US/09791537
? GENERAL INFORMATION:
? APPLICANT: Biomomix, Inc.
? APPLICANT: Debe, Derek
? APPLICANT: Danzer, Joseph
? TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
? TITLE OF INVENTION: METHODS OF USE THEREOF
? FILE REFERENCE: 261/210
? CURRENT APPLICATION NUMBER: US/09/791,537
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 153055
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 36304
? LENGTH: 628
? TYPE: PRT
? ORGANISM: Mus musculus
? US-09-791-537-36304

```

Query Match	84.3%;	Score 2734;	DB 21;	Length 628;
Best Local Similarity	84.4%;	Pred. No. 2e-220;		
Matches 534;	Conservative 47;	Mismatches 46;	Indels 6;	Gaps 3;

```
QY 1 MDTMMNLVNRILFEOLVRREVEILLSEGN-VQFQTLAKDEDFEKKKQRTIDELGKYYKDLIM 59
   ||| |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MDTMMNLVNRILFEOLVRREVEILLSEGNESIEFIQVAKDEDFEKKKQRTQNOLEKEFKDILL 60
   60 KAETERSLDVYKLLHARNOVVEIKRRQPAEDCKLEQIOLIREMIMCDTMSGSQISE 119
```

D20	61	KAFTGSAIDVAKTKARNNOVYEITRRQBAHECAKLEQOJQLRLDLMCDTSSIOJSE	120
QY	120	EOKSALAFINRGQPSSNAGNKRJSTIDESGJISLSDISFDTKDESLDMDSLVKTEFLK	179
D20	121	EOKSALAFINRGQASSGHAGNRRJSTIDESGJISLSDISFDTKDESLDMDSLVKNFMKK	180
QY	180	REKRBSSTQFVDSGPPGPKYKTRJSGSAVDGNESTYAKTVYVPNDGPIEASTIETV	239
D20	181	REKRBSSTQFIDGPPGPKYKTCJSGSVDDANESYAKTVYVPDSGGEIYASTIETL	240
QY	240	PYTRSRRTKGTLOPWNBSDTJNSKQLEPRTETDSVGTPOSGNGMRJLHDFVSKVIKES	299
D20	241	PSMTRSRGKSGPGLQPNSSALNSRPLEPRDITDMLGTPQNTGGMRLHDFVSKVIKES	300
QY	300	CVPQCKRIRFKGLJSLKCRDQRYVSHPEBCRDRCPLCPITFLIGPVYKIGEGMLADPVSQTS	359
D20	301	CVPQCKRIRFKGLJSLKCRDQRYVSHPEBCRDRCPLCPITFLIGPVYKIGEGMLADPVSQAS	360
QY	360	PMIPSIYVHCVAIEIDRGJTEGTGLRISGCRPYVEJKEKFLYKTVPLLSKVDJHAIIC	419
D20	361	PMIPSIYVHCVAIEIDRGJTEGTGLRISGCRPYVEJKEKFLYKTVPLLSKVDJHAIIC	420
QY	420	SLJLKDQFLNLKEPILLTFLNRAFMBAETDEDSIAAMYQAVGELPOANRDTLAFJLMIH	479
D20	421	SLJLKDQFLNLKEPILLTFLNRAFMBAETDEDSIAAMYQAVGELPOANRDTLAFJLMIH	480
QY	480	LQWVQASPTKMDVANIKAKEGPTTVAAVAVNPPVPMYSQOIKRQPVYRRLSLPLEYV	539
D20	481	LQWVQASPTKMDVANIKAKEGPTTVAAVAVNPPVPMYSQOIKRQPVYRRLSLPLEYV	540
QY	540	SOQWMAQENIDNIPJLHVIENTSNAFSTPQPPDIKIVSLGJFVTPPEHQJLKTSSSSLSQVR	599
D20	541	NQPMVAVDQENIDS----QKNGNNSIPRPHDVKVSLGJFVTPPEHQJLKTSSLSQVR	596
QY	600	STJTKTNPFGSKSKSATNLGQGNFPAJPMK	632
D20	597	N--LSKSTPFRGKSKSATNLGQGNFPAJPMK	628

```

RESULT 9
US-09-881-736-4
Sequence 4, Application US/09881736
GENERAL INFORMATION:
  APPLICANT: Giotzer, Michael
  APPLICANT: Jantsch-Plunger, Verena
  APPLICANT: Romano, Alper
  APPLICANT: Mishima, Masanori
  APPLICANT: Kalina, Susanne
  TITLE OF INVENTION: Cyt-4 polypeptides, DNA molecules encoding them and their use
  TITLE OF INVENTION: screening methods
  FILE REFERENCE: 0652.226001/EKS/AES
  CURRENT APPLICATION NUMBER: US/09/881,736
  CURRENT FILING DATE: 2001-06-18
  PRIOR APPLICATION NUMBER: EP 00 112 880.0
  PRIOR FILING DATE: 2000-06-19
  PRIOR APPLICATION NUMBER: EP 01 110 554.1
  PRIOR FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/241,231
  PRIOR FILING DATE: 2000-10-18
  PRIOR APPLICATION NUMBER: To be determined
  PRIOR FILING DATE: 2001-06-13
  NUMBER OF SEQ ID NOS: 6
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 4
  LENGTH: 628
  TYPE: PRT
  ORGANISM: Mus musculus
US-09-881-736-4

```

Query Match	84.3%;	Score 2734;	DB 22;	Length 628;
Best Local Similarity	84.4%;	Pred. No. 2e-220;		
Matches 534;	Conservative 47;	Mismatches 46;	Indels 6;	Gaps 3;



Db 140 MLDIKRQPKVVERLLSLPLEYWSQFMVVEQENIDPLHVIENSNAFSTPQTDIKVSLG 199  
 QY 577 PVTPEHQLKTPSSSSLSQVRSTLTPTKTPRFGSKSKATNLGRGNFASPMLK 632  
 Db 200 PVTPEHQLKTPSSSSLSQVRSTLTPTKTPRFGSKSKATNLGRGNFASPMLK 255

RESULT 12  
 US-60-412-418-1045

; Sequence 1045, Application US/60412418  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Bojn D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Wainock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; FILE REFERENCE: 660088, 465P3  
 ; CURRENT APPLICATION NUMBER: US/60/412,418  
 ; CURRENT FILING DATE: 2002-09-20  
 ; NUMBER OF SEQ ID NOS: 3025  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1045  
 ; LENGTH: 255  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-60-412-418-1045

Query Match 36.4%; Score 1182; DB 27; Length 255;  
 Best Local Similarity 99.2%; Pred. No. 2.3e-90;  
 Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 397 KEKFLRVKTVPLSKYVDIDHAICSLKDFLRNLKEPLLFRLNRAFMEEAITDEDNISIA 456  
 Db 20 RLFKFLRVKTVPLSKYVDIDHAICSLKDFLRNLKEPLLFRLNRAFMEEAITDEDNISIA 79  
 QY 457 AMYQAVGELPQANRDTLAFMLHIDQVAVOSPTKMDVANKAVFGFTYAAHAVPNDDPYT 516  
 Db 80 AMYQAVGELPQANRDTLAFMLHIDQVAVOSPTKMDVANKAVFGFTYAAHAVPNDDPYT 139  
 QY 517 MSODIKRQPKVVERLLSLPLEYWSQFMVVEQENIDPLHVIENSNAFSTPQTDIKVSLG 576  
 Db 140 MLDIKRQPKVVERLLSLPLEYWSQFMVVEQENIDPLHVIENSNAFSTPQTDIKVSLG 199  
 QY 577 PVTPEHQLKTPSSSSLSQVRSTLTPTKTPRFGSKSKATNLGRGNFASPMLK 632  
 Db 200 PVTPEHQLKTPSSSSLSQVRSTLTPTKTPRFGSKSKATNLGRGNFASPMLK 255

RESULT 13

US-09-758-472-6751  
 ; Sequence 6751, Application US/09758472  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PH001  
 ; CURRENT APPLICATION NUMBER: US/09/758,472  
 ; CURRENT FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/179,065  
 ; PRIOR FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: 60/180,628  
 ; PRIOR FILING DATE: 2000-02-04  
 ; NUMBER OF SEQ ID NOS: 9632  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6751  
 ; LENGTH: 243  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE

; LOCATION: (233)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (238)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (240)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (242)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-758-472-6751

Query Match 34.0%; Score 1103.5; DB 21; Length 243;  
 Best Local Similarity 97.4%; Pred. No. 8.5e-84;  
 Matches 222; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDTMLNVRNLFQOLVRYEILSEGNEVOFIOLAKDFEDFRKKWQRTDHELGRYKDLMLK 60  
 Db 4 MDTMLNVRNLFQOLVRYEILSEGNEVOFIOLAKDFEDFRKKWQRTDHELGRYKDLMLK 63  
 QY 61 AETERSALDVYKLNHARNOVDVEIKRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120  
 Db 64 AETERSALDVYKLNHARNOVDVEIKRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 123  
 QY 121 OKSALAFNLRGQSSNNAKNRLSTIDEGSILSDISFPKTDSDLDWSSLYKTFYLRKR 180  
 Db 124 OKSALAFNLRGQSSNNAKNRLSTIDEGSILSDISFPKTDSDLDWSSLYKTFYLRKR 183  
 QY 181 EKRRSTROFVGGPPGVKTRTSGSAVQGNESIYAKTTVYVNDGG 228  
 Db 184 EKRRSTROFVGGPPGVKTRTSGSAVQGNESIYAKTTVYVNDGG 230

RESULT 14

US-10-235-926-6751  
 ; Sequence 6751, Application US/10235926  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PH001CIN  
 ; CURRENT APPLICATION NUMBER: US/10/235,926  
 ; CURRENT FILING DATE: 2002-09-06  
 ; PRIOR APPLICATION NUMBER: 09/758,472  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: 60/179,065  
 ; PRIOR FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: 60/180,628  
 ; PRIOR FILING DATE: 2000-02-04  
 ; NUMBER OF SEQ ID NOS: 9632  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6751  
 ; LENGTH: 243  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; LOCATION: (233)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; LOCATION: (238)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; LOCATION: (240)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; LOCATION: (242)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-10-235-926-6751

Query Match 34.0%; Score 1103.5; DB 1; Length 243;  
Best Local Similarity 97.4%; Pred. No. 8.5e-84;  
Matches 222; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 MOTMMLNVRLEFQLYRRVILSEGNVOTIOLAKDFEDFRKKQRTDHELKYYKDLMLK 60

Db 4 MOTMMLNVRLEFQLYRRVILSEGNVOTIOLAKDFEDFRKKQRTDHELKYYKDLMLK 63

QY 61 AETERALDVKLKHARNOVVEIKRORAPADECKLEKROLOIREMLMCDTSGSIOLSEE 120

Db 64 AETERALDVKLKHARNOVVEIKRORAPADECKLEKROLOIREMLMCDTSGSIOLSEE 123

QY 121 OKSALAFNLNGOPSSNAGNKRSLTIDSGSILSDISFDKTDSDLDWSSLVYTFKLLKR 180

Db 124 OKSALAFNLNGOPSSNAGNKRSLTIDSGSILSDISFDKTDSDLDWSSLVYTFKLLKR 183

QY 181 EKRRSTSRQFVDPGPPGVKTRISGSAVDGNSIYAKTIVYVNDGG 228

Db 184 EKRRSTSRQFVDPGPPGVKTRISGSAVDGNSIYAKTIVYVNDGG 230

## RESULT 15

PCT-US01-08656-9072

; Sequence 9072, Application PC/TUS0108656

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 21272-066

; CURRENT APPLICATION NUMBER: PCT/US01/08656

; CURRENT FILING DATE: 2001-04-16

; PRIOR APPLICATION NUMBER: 09/522,929

; PRIOR FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: 09/770,160

; PRIOR FILING DATE: 2001-01-26

; NUMBER OF SEQ ID NOS: 10994

; SOFTWARE: Custom

; SEQ ID NO 9072

; LENGTH: 256

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (7)..(33)

; OTHER INFORMATION: PROTEIN GTPASE DOMAIN ACTIVATION domain identified by

; OTHER INFORMATION: EMATRIX, accession number PD00930A, p-value=7.324e-15, raw score

; CT-US01-08656-9072

Query Match 30.0%; Score 971.5; DB 1; Length 256;

Best Local Similarity 84.0%; Pred. No: 1.2e-72;

Matches 205; Conservative 3; Mismatches 27; Indels 9; Gaps 4;

QY 357 QTSPIPISTIVHGVNIEQGLTETGIRISGCDRTYKELKEFLRYKYTPLLSKYDDIH 416

Db 1 QTSPIPISTIVHGVNIEQGLTETGIRISGCDRTYKELKEFLRYKYTPLLSKYDDIH 60

QY 417 AICSLKDFLRLNKEPLTFRLNRAFMEAEITDEDNSIAATQAVGELPOANDTLAFL 476

Db 61 AICSLKDFLRLNKEPLTFRLNRAFMEAEITDEDNSIAATQAVGELPOANDTLAFL 120

QY 477 MHLQVVAOSPHTKMDVANLAKVGPITYAHAVPNPDPTMSQDIR-QPKVVERLLSLP 535

Db 121 MHLQVVAOSPHTKMDVANLAKVGPITYAHAVPNPDPTMSQDIR-QPKVVERLLSLP 180

QY 536 L-EYNSQFMVVEE---NIDPLAVIENSNAFSTPQTPDIKV---SLGCVTTPEHQLK 587

Db 181 LMEYWEFSFMDGEGTDPDLAVIENSNAFSTPQTPDIKAVPGGGLCVHTAGAEAIQK 240

QY 588 TPSS 591

Db 241 GPPS 244

Search completed: July 3, 2003, 10:02:52  
Job time : 307 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:50:02 ; Search time 88 Seconds  
(without alignments)  
1479.793 Million cell updates/sec

Title: US-09-881-736-2  
Perfect score: 3243  
Sequence: 1 MDTWMLNVRNLFQGLVRRVE.....SKSATNIGRCGNFSPMLK 632

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protist:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rviro:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3243	100.0	632	4	Q9P2W2
2	3237	99.8	632	4	Q9H0H5
3	3230	99.6	632	4	Q9NMN2
4	2919	90.0	570	4	Q9P250
5	2781	85.8	628	4	Q9B274
6	2734	84.3	628	11	Q9WYU1
7	1182	36.4	255	4	Q9H9L9
8	864	26.6	625	5	Q9V607
9	862	26.6	625	5	Q9N929
10	547.5	16.9	681	5	Q9XUS9
11	431	13.3	354	5	Q95VR1
12	328	10.1	334	11	Q91V57
13	322	9.9	459	4	Q96F50
14	320	9.9	459	13	Q92153
15	308	9.5	695	3	Q94466
16	305.5	9.4	1388	5	Q8ST20

17	305.5	9.4	1439	5	Q9TY64	Q9TY64 caenorhabdi
18	301	9.3	1261	4	Q15463	Q15463 mus musculu
19	294	9.1	282	11	Q9D9W2	Q9D9W2 mus musculu
20	291	9.0	970	4	Q9P107	Q9P107 mus musculu
21	290.5	9.0	733	4	Q9BYS7	Q9BYS7 mus musculu
22	290.5	9.0	759	4	Q9UJ00	Q9UJ00 mus musculu
23	290.5	9.0	814	4	Q9UNM1	Q9UNM1 mus musculu
24	287.5	8.9	263	11	Q9EP57	Q9EP57 mus musculu
25	286.5	8.8	1286	4	Q9P2C3	Q9P2C3 mus musculu
26	285.5	8.8	1165	4	Q92619	Q92619 mus musculu
27	281.5	8.7	889	4	Q12844	Q12844 mus musculu
28	281.5	8.7	936	4	Q12843	Q12843 mus musculu
29	275.5	8.5	210	11	Q9D9B3	Q9D9B3 mus musculu
30	270.5	8.3	568	13	Q98935	Q98935 gallus gall
31	270.5	8.3	1397	4	Q9NTG2	Q9NTG2 mus musculu
32	270.5	8.3	2548	4	Q9UNJ2	Q9UNJ2 mus musculu
33	269.5	8.3	735	4	Q9P227	Q9P227 mus musculu
34	267.5	8.2	2626	11	Q921N3	Q921N3 mus musculu
35	266.5	8.2	523	11	Q9CU46	Q9CU46 mus musculu
36	265	8.2	494	5	Q9W4A9	Q9W4A9 drosophila
37	260.5	8.0	802	4	Q60890	Q60890 mus musculu
38	258.5	8.0	714	3	Q13384	Q13384 mus musculu
39	258.5	8.0	786	11	Q99MT3	Q99MT3 mus musculu
40	258	8.0	903	4	Q14560	Q14560 mus musculu
41	256.5	7.9	632	4	Q96EV3	Q96EV3 mus musculu
42	256.5	7.9	714	3	Q13390	Q13390 mus musculu
43	256.5	7.9	786	4	Q96S75	Q96S75 mus musculu
44	255.5	7.9	245	11	Q99LW5	Q99LW5 mus musculu
45	254	7.8	683	11	Q99MT4	Q99MT4 mus musculu

## ALIGNMENTS

RESULT 1  
Q9P2W2 PRELIMINARY; PRT; 632 AA.

AC Q9P2W2:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE GTPase activating protein.  
 GN MGCRCACGAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=20435340; PubMed=10979956;  
 RA Kawashima T., Hirose K., Saton T., Kaneko A., Ikeda Y., Kaziro Y.,  
 RT "MGCRCACGAP is involved in the control of growth and differentiation of  
 RT hematopoietic cells.";  
 RL Blood 96:2116-2124(2000).  
 DR EMBL; AB030251; BAA80247.1; -  
 DR HSRP; Q07960; IRGP.  
 DR InterPro; IPR002219; DAG\_PE\_bind.  
 DR InterPro; IPR000198; RhogAP.  
 DR Pfam; PF00130; DAG\_PE\_bind. 1.  
 DR Pfam; PF00620; RhogAP. 1.  
 DR SMART; SM00109; CL. 1.  
 DR SMART; SM00324; RhogAP. 1.  
 DR PROSITE; PS00479; DAG\_PE\_BIND\_DOM.1; UNKNOWN.1.  
 DR PROSITE; PS50081; DAG\_PE\_BIND\_DOM.2; 1.  
 SQ SEQUENCE 632 AA; 71000 MW; BC2B75E5A8739E2B CRC64;

Query Match 100.0%; Score 3243; DB 4; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-213; Indels 0; Gaps 0;  
 Matches 632; Conservative 0; Mismatches 0;  
 1 MDTWMLNVRNLFQGLVRRVEITSGNEVQFQLKKDFEDFRKRNQRTHELGKYLDMK 60  
 |||

```

Db      1 MDTMMLNVRNLFEOQVRAVEILISEGNEVOFIOLANDEFEDFRKKWQRTDHELKGYDLMK 60
QY      61 AETERSALDVKLKHAHQVDVEIKRQRAADCEKLEKRIQIILRELMKCDTSGSIQISEE 120
        |||
Db      61 AETERSALDVKLKHAHQVDVEIKRQRAADCEKLEKRIQIILRELMKCDTSGSIQISEE 120
QY      121 OKSALAFNLKRGPPSSNAGNKRSLSTIDESGSIISDIFDKTDESILDMDSLVKTFKLRK 180
        |||
Db      121 OKSALAFNLKRGPPSSNAGNKRSLSTIDESGSIISDIFDKTDESILDMDSLVKTFKLRK 180
QY      181 EKRSTSRQFVDPGPPGVKRTSRISGSAVDQGNESIYAKTVTVPNDCGPIEAVSTIETVP 240
        |||
Db      181 EKRSTSRQFVDPGPPGVKRTSRISGSAVDQGNESIYAKTVTVPNDCGPIEAVSTIETVP 240
QY      241 YWTRSRKKTGTLOPNMSDSTLNSROLEPRTEDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
        |||
Db      241 YWTRSRKKTGTLOPNMSDSTLNSROLEPRTEDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
QY      301 VECGRKIKFGKLSLKCRCQDVVSHPECRCRCPICPTLIGTPVIGEGMLADVFVSQTSR 360
        |||
Db      301 VECGRKIKFGKLSLKCRCQDVVSHPECRCRCPICPTLIGTPVIGEGMLADVFVSQTSR 360
QY      361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKFLRYKTVPLLSKYDDIHAICS 420
        |||
Db      361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKFLRYKTVPLLSKYDDIHAICS 420
QY      421 LKDFLRNKEPDLTFFRLNRAFMEEAEITDENSTIAAMYQAVGELPOANDTLAFIMHL 480
        |||
Db      421 LKDFLRNKEPDLTFFRLNRAFMEEAEITDENSTIAAMYQAVGELPOANDTLAFIMHL 480
QY      481 ORVAOSPHTKMVDANLAKVFPTIYAAVNPDPVTMSODIKRQKVERLLSLPLEYWS 540
        |||
Db      481 ORVAOSPHTKMVDANLAKVFPTIYAAVNPDPVTMSODIKRQKVERLLSLPLEYWS 540
QY      541 QPMWVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPTVTPPHOLKTPSSSSLSQRYWS 600
        |||
Db      541 QPMWVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPTVTPPHOLKTPSSSSLSQRYWS 600
QY      601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632
        |||
Db      601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632
        |||
RESULT 2
Q9H0H5 PRELIMINARY; PRT; 632 AA.
AC Q9H0H5:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 71.0 kDa protein.
DE DKFZP434C011.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgorge W., Boecker M., Blocker H., Bauesachs S., Blum H.,
RA Lauber J., Dusterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Mambrot R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435(2001).
DR EMBL; AL136794; CAB66728.1;
DR HSSP; 007960; IRGP.
DR InterPro; IPR002219; DAG_PE-bind.
DR Pfam; PF00130; DAG_PE-bind; 1.

```

```

DR Pfam; PF00620; RhogAP. 1.
DR SMART; SM00109; C1; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 632 AA; 71026 MW; 032B7DF9CEA8F39D CRC64;
Query Match 99.8%; Score 3237; DB 4; Length 632;
Best local similarity 99.8%; Pred. No. 46-213;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 MDTMMLNVRNLFEOQVRAVEILISEGNEVOFIOLANDEFEDFRKKWQRTDHELKGYDLMK 60
        |||
Db      1 MDTMMLNVRNLFEOQVRAVEILISEGNEVOFIOLANDEFEDFRKKWQRTDHELKGYDLMK 60
QY      61 AETERSALDVKLKHAHQVDVEIKRQRAADCEKLEKRIQIILRELMKCDTSGSIQISEE 120
        |||
Db      61 AETERSALDVKLKHAHQVDVEIKRQRAADCEKLEKRIQIILRELMKCDTSGSIQISEE 120
QY      121 OKSALAFNLKRGPPSSNAGNKRSLSTIDESGSIISDIFDKTDESILDMDSLVKTFKLRK 180
        |||
Db      121 OKSALAFNLKRGPPSSNAGNKRSLSTIDESGSIISDIFDKTDESILDMDSLVKTFKLRK 180
QY      181 EKRSTSRQFVDPGPPGVKRTSRISGSAVDQGNESIYAKTVTVPNDCGPIEAVSTIETVP 240
        |||
Db      181 EKRSTSRQFVDPGPPGVKRTSRISGSAVDQGNESIYAKTVTVPNDCGPIEAVSTIETVP 240
QY      241 YWTRSRKKTGTLOPNMSDSTLNSROLEPRTEDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
        |||
Db      241 YWTRSRKKTGTLOPNMSDSTLNSROLEPRTEDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
QY      301 VECGRKIKFGKLSLKCRCQDVVSHPECRCRCPICPTLIGTPVIGEGMLADVFVSQTSR 360
        |||
Db      301 VECGRKIKFGKLSLKCRCQDVVSHPECRCRCPICPTLIGTPVIGEGMLADVFVSQTSR 360
QY      361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKFLRYKTVPLLSKYDDIHAICS 420
        |||
Db      361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKFLRYKTVPLLSKYDDIHAICS 420
QY      421 LKDFLRNKEPDLTFFRLNRAFMEEAEITDENSTIAAMYQAVGELPOANDTLAFIMHL 480
        |||
Db      421 LKDFLRNKEPDLTFFRLNRAFMEEAEITDENSTIAAMYQAVGELPOANDTLAFIMHL 480
QY      481 ORVAOSPHTKMVDANLAKVFPTIYAAVNPDPVTMSODIKRQKVERLLSLPLEYWS 540
        |||
Db      481 ORVAOSPHTKMVDANLAKVFPTIYAAVNPDPVTMSODIKRQKVERLLSLPLEYWS 540
QY      541 QPMWVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPTVTPPHOLKTPSSSSLSQRYWS 600
        |||
Db      541 QPMWVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPTVTPPHOLKTPSSSSLSQRYWS 600
QY      601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632
        |||
Db      601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632
        |||
RESULT 3
Q9NMN2 PRELIMINARY; PRT; 632 AA.
AC Q9NMN2:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CDNA FLJ20726 fis. clone HEP13735.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Odayashi M.,

```

RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
RT "NEO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK000733; BAA91347.1; -  
DR HSSP: Q07960; IRGP.  
DR InterPro: IPR002219; DAG\_PE-bind.  
DR InterPro: IPR000198; RhogAP.  
DR Pfam: PF00130; DAG\_PE-bind; 1.  
DR Pfam: PF00620; RhogAP; 1.  
DR SMART: SM00109; Cl; 1.  
DR SMART: SM00324; RhogAP; 1.  
DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; UNKNOWN\_1.  
DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM\_2; 1.  
SQ SEQUENCE 632 AA; 71048 MW; 7DC06CB725BBA4A CRC64;

Query Match 99.68; Score 3230; DB 4; Length 632;  
Best Local Similarity 99.78; Pred. No. 1,2e-212;  
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MOTMMLNVRNLFQOLVRVEIISSEGNVDFIOLAKDFEDRRKKMORTDHELGKYLK 60  
DB 1 MOTMMLNVRNLFQOLVRVEIISSEGNVDFIOLAKDFEDRRKKMORTDHELGKYLK 60  
QY 61 AETERSALDVKLKHAHQVDVEIKRQRAEADCEKLERQIOLIRELMCDTSGSIOLSEE 120  
DB 61 AETERSALDVKLKHAHQVDVEIKRQRAEADCEKLERQIOLIRELMCDTSGSIOLSEE 120  
QY 121 QKSALFLNRGQSSSNAGNKRSLSTIDEGSILSDISFQKTDSDLDMSLVYTFIKKR 180  
DB 121 QKSALFLNRGQSSSNAGNKRSLSTIDEGSILSDISFQKTDSDLDMSLVYTFIKKR 180  
QY 181 EKRSTSRQFVDPGPGPVKTRISGSAVDGNEISYAKTVVVPNDGPIEAVSTIETVP 240  
DB 181 EKRSTSRQFVDPGPGPVKTRISGSAVDGNEISYAKTVVVPNDGPIEAVSTIETVP 240  
QY 241 YTRSRRTGTLOPWNDSSTLNSRQLEPRTETDSVGTPOSGMRLHDFVSKTVIKPESC 300  
DB 241 YTRSRRTGTLOPWNDSSTLNSRQLEPRTETDSVGTPOSGMRLHDFVSKTVIKPESC 300  
QY 301 VPCKRIKFGKLSLKCRCRVVSHPECRDRCPLCIPILIGTPVKIGEGMLADFVSQTS 360  
DB 301 VPCKRIKFGKLSLKCRCRVVSHPECRDRCPLCIPILIGTPVKIGEGMLADFVSQTS 360  
QY 361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKELRYKTVPLSKVDIHAICS 420  
DB 361 MIPSIYVHCVNEIEQGLTETGLYRISGCDRTYKELKEKELRYKTVPLSKVDIHAICS 420  
QY 421 LKQDLRLNKEPILFRLNRAFEAAEITDEDSIAAMYQAVGELPOANDTLAFIMHL 480  
DB 421 LKQDLRLNKEPILFRLNRAFEAAEITDEDSIAAMYQAVGELPOANDTLAFIMHL 480  
QY 481 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
DB 481 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
QY 541 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
DB 541 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
QY 601 TLTKNTPRFGSKSKSATNLGRGNEFPASPMLK 632  
DB 601 TLTKNTPRFGSKSKSATNLGRGNEFPASPMLK 632  
RESULT 4  
Q9P250 PRELIMINARY; PRT; 570 AA.  
AC Q9P250;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE KIAA1478 protein (Fragment).  
GN KIAA1478.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2077482; PubMed=10819331;  
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human  
genes. XVII. The complete sequences of 100 new cDNA clones from brain  
RT which code for large proteins in vitro."  
RL DNA Res. 7:143-150(2000).  
DR EMBL: AB040911; BAA96002.1; -  
DR HSSP: Q07960; IRGP.  
DR InterPro: IPR002219; DAG\_PE-bind.  
DR InterPro: IPR000198; RhogAP.  
DR Pfam: PF00130; DAG\_PE-bind; 1.  
DR Pfam: PF00620; RhogAP; 1.  
DR SMART: SM00109; Cl; 1.  
DR SMART: SM00324; RhogAP; 1.  
DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; UNKNOWN\_1.  
DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM\_2; 1.  
FT NON\_TER  
SQ SEQUENCE 570 AA; 63468 MW; A03F36C0AC65485 CRC64;

Query Match 99.08; Score 2919; DB 4; Length 570;  
Best Local Similarity 99.88; Pred. No. 1.9e-191;  
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 63 TERSALDVKLKHAHQVDVEIKRQRAEADCEKLERQIOLIRELMCDTSGSIOLSEE 122  
DB 1 TERSALDVKLKHAHQVDVEIKRQRAEADCEKLERQIOLIRELMCDTSGSIOLSEE 122  
QY 123 SALAFNLNRGQSSSNAGNKRSLSTIDEGSILSDISFQKTDSDLDMSLVYTFIKKR 182  
DB 123 SALAFNLNRGQSSSNAGNKRSLSTIDEGSILSDISFQKTDSDLDMSLVYTFIKKR 182  
QY 183 RSTSRQFVDPGPGPVKTRISGSAVDGNEISYAKTVVVPNDGPIEAVSTIETVP 242  
DB 183 RSTSRQFVDPGPGPVKTRISGSAVDGNEISYAKTVVVPNDGPIEAVSTIETVP 242  
QY 241 YTRSRRTGTLOPWNDSSTLNSRQLEPRTETDSVGTPOSGMRLHDFVSKTVIKPESC 300  
DB 241 YTRSRRTGTLOPWNDSSTLNSRQLEPRTETDSVGTPOSGMRLHDFVSKTVIKPESC 300  
QY 303 CGKRIKFGKLSLKCRCRVVSHPECRDRCPLCIPILIGTPVKIGEGMLADFVSQTS 362  
DB 303 CGKRIKFGKLSLKCRCRVVSHPECRDRCPLCIPILIGTPVKIGEGMLADFVSQTS 362  
QY 421 LKQDLRLNKEPILFRLNRAFEAAEITDEDSIAAMYQAVGELPOANDTLAFIMHL 480  
DB 421 LKQDLRLNKEPILFRLNRAFEAAEITDEDSIAAMYQAVGELPOANDTLAFIMHL 480  
QY 481 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
DB 481 QRYAOSPHTKMDVANLAKYGPPIYAHAVPNPDPYMSODIKRQPVVEELSLPLEYMS 540  
QY 543 MMYEÖENIDPLAHYIENSNAFSTPOTPDIKYSLGPTTPEHOLKTPSSSLSORVSTL 602  
DB 543 MMYEÖENIDPLAHYIENSNAFSTPOTPDIKYSLGPTTPEHOLKTPSSSLSORVSTL 602  
QY 603 TKNTPRFGSKSKSATNLGRGNEFPASPMLK 632  
DB 603 TKNTPRFGSKSKSATNLGRGNEFPASPMLK 632  
RESULT 5  
Q9B274 PRELIMINARY; PRT; 628 AA.  
AC Q9B274;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE KIAA1478 protein (Fragment).  
GN KIAA1478.



```

QY 180 REKRSTSRQFVDPGPPGPKTKRTSRIGSAVDGNEISVAKTQVTPNDGSPREAVSTETV 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 REKRSTSRQFIDGPPGPKTKCTSIGSTVDQANESIVAKTQVTPNDGSPREAVSTETV 240
QY 240 PWTTSRRTGTGLOPNSNSTLNSROLERTETDSGTQSGMRLHDFVSKTVKPS 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 PWTTSRRTGTGLOPNSNSTLNSROLERTETDSGTQSGMRLHDFVSKTVKPS 300
QY 300 CVPCKRRTKFGKLSKRCRCRVVSHPECRCPLPCIPILGIPVYIGGMLADPFQSS 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CVPCKRRTKFGKLSKRCRCRVVSHPECRCPLPCIPILGIPVYIGGMLADPFQSS 360
QY 360 PMIPSIYVNCVNEIDRGRLTETGLYRISGCDRTYKELKEFLRVKTVPLSLKYDDIHAIC 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 PMIPSIYVNCVNEIDRGRLTETGLYRISGCDRTYKELKEFLRVKTVPLSLKYDDIHAIC 420
QY 420 SLTKDFLRKLKPLLTFLRLNRAFMKAEITTEDNSIAAMYQAVGELPQANRDTLAFIMH 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 SLTKDFLRKLKPLLTFLRLNRAFMKAEITTEDNSIAAMYQAVGELPQANRDTLAFIMH 480
QY 480 LQRVASQSPHTKMDVANLAKVFGPTIYAAHVNPDPVTMSQDTRKQKVERLLSLPLETW 539
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 LQRVASQSPHTKMDVANLAKVFGPTIYAAHVNPDPVTMSQDTRKQKVERLLSLPLETW 540
QY 540 SQFMVVEQENIDPLHVIENSNAFSTPQTPDIKVSILGPTTPEHOLKTPSSSSLSQRYR 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 NQFMVVEQENIDPLHVIENSNAFSTPQTPDIKVSILGPTTPEHOLKTPSSSSLSQRYR 596
QY 600 SUTLNTPRFGSKSKSATNLGROGNFPAAPMLK 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 597 N-LSKSTPRFGSKSKSATNLGROGNFPAAPMLK 628

```

## RESULT 7

```

QY 09H9L9 PRELIMINARY: PRT: 255 AA.
ID 09H9L9
AC 09H9L9
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CDNA FLJ12664 f15, clone NTZRM4002226, weakly similar to GTPase
activating protein rotund.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugeno S., Shitatori A., Sudo H.,
RA Watanabe M., Hosogi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.,
RT "NEBO human cDNA sequencing project."
RT Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK023726; BAB14206.1; -
DR HSSP: Q07960; 1RCP.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR00198; RhogAP.
DR Pfam: PF00620; RhogAP.
DR SMART: SM00324; RhogAP.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN.1.
SQ SEQUENCE 255 AA; 28488 MW; B59D7B59385821BD CRC64;

```

```

QY Query Match 36.4%; Score 1182; DB 4; Length 255;
    Best Local Similarity 99.2%; Pred. No. 3.9e-73;
    Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 397 KKKFLRVKTVPLSLKYDDIHAICSLKDFLRNKLKEPLTLFLRLNRAFMKAEITTEDNSIA 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20 KKKFLRVKTVPLSLKYDDIHAICSLKDFLRNKLKEPLTLFLRLNRAFMKAEITTEDNSIA 79

```

```

QY 457 AMYQAVGELPQANRDTLAFIMHQLQVASQSPHTKMDVANLAKVFGPTIYAAHVNPDPVT 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 80 AMYQAVGELPQANRDTLAFIMHQLQVASQSPHTKMDVANLAKVFGPTIYAAHVNPDPVT 139
QY 517 MSQDTRKQKVERLLSLPLETWQFMVVEQENIDPLHVIENSNAFSTPQTPDIKVSILG 576
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 MSQDTRKQKVERLLSLPLETWQFMVVEQENIDPLHVIENSNAFSTPQTPDIKVSILG 199
QY 577 PYTTPPEHOLKTPSSSSLSQRYRSTTKTTPRFGSKSKSATNLGROGNFPAAPMLK 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 PYTTPPEHOLKTPSSSSLSQRYRSTTKTTPRFGSKSKSATNLGROGNFPAAPMLK 255

```

## RESULT 8

```

QY 09V6U7 PRELIMINARY: PRT: 625 AA.
ID 09V6U7
AC 09V6U7
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CG13345 protein.
GN RACGAP50C OR CG13345.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RC MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brotlier P.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Bolshakov S.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jaitai M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner A., Turner R., Venter A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003816; AAF58324.1; -
DR FlyBase; FBgn003381; RacGAP.
DR InterPro; IPR002219; DAG_PE-bind.

```









```

QY 259 STLNROLEPRTETDVSQVGTPOSGMRLHDFVSKTVIKPESCVPCKRIKFGKLS--LKC 316
DB 31 ATLKENEHVPEKE-----KVNFKVHFHFRGPHMCEYCA-NMMGLIAGVVC 76
QY 317 RDCRVVSHPECRDRCPLPCIPPL-----IGTPVKIGEGMLADPFVSQTSPIPISTIV 366
DB 77 ADCCGLNVHKQCSKMPNDCKPDLKHKKVYSCDLFTLVK-----AHFKR-----PMVV 125
QY 367 VHCVEIQRGLTEGLVIRISGCDRTVKELEKFLR--VKTVPLSKVDIHAICSLMD 424
DB 126 DMCIEIRISRGINSGLRVSGFSDLEOVKMAFPDGEKADISVMYEDINITGALTL 185
QY 425 FLRNKEPLLTFLRNRAFMFAEITDEDNSIAMYOAVGELPOANRDTLAFIMHLQRYA 484
DB 186 YRDLPIPLITDAPKFTIESAKTDPDQGLIHLHALLLPRAHCETLIRYLAHAKRY- 244
QY 485 OSPHKK---MDVANLAKVFGPTIVAHAVPNPDPTMSODIKRQKVERLL 532
DB 245 -TLHKEKMLMSAENIGIVFGPTLM--RAPELDAMALANDIRYQRLVEMELI 292

```

## RESULT 15

```

ID 094466 PRELIMINARY; PRT: 695 AA.
AC 094466;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN Putative GTPase-activating protein C23G7.08C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Xiang Z., Aves S., Rajandream M.A., Barrell B.G.;
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC RHO GDP-DISSOCIATION INHIBITORS.
DR EMBL; AL035065; CAA22624.1; -.
DR HSSP; Q07960; IRGP.
DR InterPro: IPR001060; Cdc15_Fes_CIP4.
DR InterPro: IPR00198; RhogAP.
DR Pfam; PF00611; FCH; 1.
DR Pfam; PF00620; RhogAP; 1.
DR SMART; SM00055; FCH; 1.
DR SMART; SM00324; RhogAP; 1.
DR Hypothetical protein; GTPase activation.
KW DOMAIN 420
FT DOMAIN 420
SQ SEQUENCE 695 AA; 76678 MW; IDEAF1328D576263 CRC64;

```

## Query Match

Best Local Similarity 9.5%; Score 308; DB 3; Length 695;

Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23;

```

QY 35 KDEFEFRKKWQRTDHE---LGYKDLKKAETE-RSALDYKLHARNQVDYEIK--RR 86
DB 145 KSIKYEYARKQENAVYLEAVWQMDKSKSRFGAETENRALDNK-----NTGDSQKQGFVK 199
QY 87 ORAEADCEKLEROIOLIREMLMCDTSGSIOLSEE-OKSALAF-----LNRGOPS 134
DB 200 PRSNAQLTKLEDEARLKAENAESDMKSIENANOVOKOLCTHRPNYIKQFSLQREIES 259
QY 135 SSNAQNKRLSTIDESGSLSDI-----SFDKTDLSLDDSSLVKTFKLK 178
DB 260 SLIANYLRATKLCESNTLLNGLTIRPKPTPTNGCIGRHLNDINANTDPEVQVLAHA-SIK 318
QY 179 KREKRKRST--SRQFVDPGPPGPKVTKTRISGSAVDGNESTIAKTYTVTPNDGGPIE---A 232
DB 319 HEDNNRPDASKTKIIOPPSSYGTGSSAGKTNPNVNPFI--KVTAAPIS---PLQNTNPA 373

```

```

QY 233 VSTIETVPYWTRSRRTGTILQPMNSDSTLN--SROLEPRTETDVSQVGTPOSGMRLHDFV 290
DB 374 PSTFPNPVSASPAFPNNSSTNSPSTAPASAPLASTIKPSTANDTNGSSSSSS----- 426
QY 291 SKTVIKPESCVPCKRIKFGKLSLKCRCRVVSHPECRDRCPLPCIP-----TLIGT- 342
DB 427 -----NPRTPSPPLASNAE-----NKPPVQAQSPVLLPLPIQTITIGTS 467
QY 343 -----PVKIGEGMLADPFVSQTS-----PMIPSTIVHCY 370
DB 468 REVAPPSSINSNRASPPRPTSVSPQSPPTKSLFGARLDAITLREHNSNPNIWMOCT 527
QY 371 NEIEQGLTEGLVIRISGCDRTVKELEKFLRVKTVPL-----SKVDIHAICSLMD 424
DB 528 SOYENGLNDGITYRVPSSARVNMILRSQF---ENNPLQLHTPEDEYDVAHVADLLKI 584
QY 425 FLRNKEPLLTFLRNRAFMFAEITDEDNSIAMYOAVGELPOANRDTLAFIMHLQRYA 484
DB 585 FPRELREPLIPNHQRDFIDAGNVEDESRRDVAHRAINDLPANYSTIRHLTHLAKIK 644
QY 485 Q-SPHTKMDVANLAKVFGPTIVAHAVPNPDPTMSODIK 522
DB 645 ENSDVNKMSTNNLAIIMGPTIIRKQAT-IEIISFSFTIE 682

```

Search completed: July 3, 2003, 09:56:17

Job time : 92 secs





XX	(HXSE-) HXSEQ INC.
PA	
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Pt	Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI; 2001-442253/47.
XX	N-PSDB; AAI59203.
PT	
Pr	Novel nucleic acids and polypeptides, useful for treating disorders
XX	such as central nervous system injuries -
PS	
XX	Example 4; SEQ ID NO 3192; 10078bp; English.
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM8642-AAM42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilitation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
XX	Sequence 632 AA:
SQ	
Query Match	99.7%; Score 3233; DB 22; Length 632;
Best Local Similarity	99.7%; Pred. No. 6.8e-263;
Matches 630; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 MDTMMANRNLFEOLYRVEVLLISENEVOFIOLANDFEDFRKKMORTHELKGYDLMK 60
Db	1 MDTMALNRLNFEOLYRREVELLISENEVOFIOLANDFDFRRKMORTHELEKYDLMK 60
OY	AETERSALDVKLKHARNOVDVEIKRKORAEADCEKLEIOILIRFMKMDTSGSIQLSBE 120
Db	61 AETERSALDVKLKHARNVDVEIKRKORAEADCEKLEIQILIREMKMDTSGSIQLSBE 120
OY	121 OKSALAFLNRQPSSNAGNKRRLSTIDEGSILDSISFDKTDESLDWSSLVTKFKLR 180
Db	121 OKSALAFLNRQPSSNAGNKRRLSTIDEGSILDSISFDKTDESDMWSSLVTKFKLR 180
OY	181 EKRRTSQPFDPGPGPYKPKTRISGSANDOGNESIVATTVYVPNDGGPIEAVSYIEHP 240
Db	181 EKRRTSQPFDPGPGPYKPKTRISGSANDOGNESIVATTVYVPNDGGPIEAVSYIEHP 240
OY	241 TWRSRRKGTGLQPNWSDTLSNSRQLBERTETDSVGTFQPSNGMRLHDFSVKTIKIPESC 300
Db	241 TWRSRRKGTGLQPNWSDTLSNSRQLBERTETDSVGTFQPSNGMRLHDFSVKTIKIPESC 300
OY	301 VPCKRIKEGLSLKRCDCRVVSHPECDRCPLPICITPLIGPVKIGGMADPVSQSFP 360
Db	301 VPCKRIKEGLSLKRCDCRVVSHPECDRCPLPICITPLIGPVKIGGMADPVSQSFP 360
OY	361 MIPSIVACVNEIEDRGITGLVIRISGCDTVKELKFKLVKVPPLISKVDTHAICS 420
Db	361 MIPSIVACVNEIEDRGITGLVIRISGCDTVKELKFKLVKVPPLISKVDTHAICS 420
OY	421 LKDFLNRLKPELLTLRLNRAFMEEAITDEDNSTIAAMYQAVGELPDQNRDTLAFLMTHL 480
Db	421 LKDFLNRLKPELLTLRLNRAFMEEAITDEDNSTIAAMYQAVGELPDQNRDTLAFLMTHL 480
OY	481 ORVASOPTKADVANIKNVGPVTVAHVAVNPDPVTMSODIKRKQKVVERLISTPLEWS 540
Db	481 ORVASOPTKADVANIKNVGPVTVAHVAVNPDPVTMSODIKRKQKVVERLISTPLEWS 540

Qy	54	QEWVWVQENIDPLHVENSNNAFTQTPDIKXSLGLGPTTPEHOLKTPSSSLSSQVR	600
Db	541	QFWWVQENIDPLHVENSNNAFTQTPDIKXSLGLGPTTPEHOLKTPSSSLSSQVR	6000
Qy	601	TLTKNTPRFGSKSKSATNLGRQGNFPASPMLK	632
Db	601	TLTKNTPRFGSKSKSATNLGRQGNFPASPMLK	632
RESULT 2			
ID	AAU69419	standard; Protein; 632 AA.	
XX	AAU69419;		
AC	AAU69419;		
XX	30-JAN-2002 (first entry)		
DT			
XX	Lung small cell carcinoma antigen #13.		
DE			
XX	Human; cytostatic; antitumour; lung small cell cancer antigen;		
KW	tumour; lung cancer.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200177168-A2.		
PD	18-OCT-2001.		
XX			
PF	11-APR-2001; 2001WO-US11859.		
XX			
PR	11-APR-2000; 2000US-196780P.		
PR	21-JUN-2000; 2000US-213361P.		
PR	01-SEP-2000; 2000US-229763P.		
PR	05-SEP-2000; 2000US-230629P.		
PR	14-SEP-2000; 2000US-232656P.		
PR	19-DEC-2000; 2000US-257037P.		
PR	08-JAN-2001; 2001US-260796P.		
PA	(CORI-) CORIXA CORP.		
XX			
FI	Lodes MJ, Wang T, Mohamath R, Indirias CT;		
XX			
DR	WPI: 2002-010896/01.		
DR	N-PSDB; AAS61859.		
PT	Lung tumour polynucleotide and polypeptides useful in therapy and		
XX	diagnosis of cancer especially lung cancer		
XX			
XX	Claim 2; Page 270-271; 295pp; English.		
XX			
CC	The invention relates to novel isolated lung small cell cancer antigen		
CC	polynucleotides (I) and polypeptides (II) used in a method of detecting		
CC	cancer in a patient. The method is optionally performed by		
CC	utilising oligonucleotides (III), where the biological sample		
CC	from the patient is contacted with (III), detecting the amount of		
CC	polynucleotide hybridised to (III) in the sample and comparing the		
CC	amount of polynucleotide to a predetermined cut-off value and thereby		
CC	determining cancer in a patient. (I), (II) or antigen-presenting cells		
CC	expressing (II) is useful for stimulating and/or expanding T cells		
CC	specific for a tumour protein. The method comprises contacting T cells		
CC	with one of the components under conditions to permit the stimulation		
CC	and/or expansion of the cells. A composition comprising (I) is useful for		
CC	stimulating an immune response in a patient and for inhibiting the		
CC	development of a cancer especially lung cancer in a patient. An		
CC	isolated T cell population is useful for removing tumour cells from the		
CC	biological sample and for inhibiting the development of cancer in a		
CC	patient. AAU69407-AAU69431 represent novel human lung small cell		
CC	cancer antigen amino acid sequences of the invention.		
XX			
SQ	Sequence	632 AA;	

Query Match
99.6%; Score 3230; DB 23; Length 632;

Best Local Similarity 99.7%; Pred. No. 1.2e-262;  
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MDTMNLNVLNLEFQOLYRVEILISEGNEVFIOLAKFEDEPRKKMOTDEHLKYKLLMK 60
DB 1 MDTMNLNVLNLEFQOLYRVEILISEGNEVFIOLAKFEDEPRKKMOTDEHLKYKLLMK 60
OY 61 AETERSALDVKLKHARNOVDVEIKRQRAEADCEKLEROIOLIREMLMCDTSGSIOLSE 120
DB 61 AETERSALDVKLKHARNOVDVEIKRQRAEADCEKLEROIOLIREMLMCDTSGSIOLSE 120
OY 121 OKSALAFLNRGOPSSSSNAGNKRLLSTIDSGSILSDISFDKTESLDMDSSLVYKTFKKR 180
DB 121 OKSALAFLNRGOPSSSSNAGNKRLLSTIDSGSILSHISFDKTESLDMDSSLVYKTFKKR 180
OY 181 EKRSTSRQFVGGPPPVKKTSTISGAVDQGNESIYAKTIVYVPPNDGPIEAVSTIETVP 240
DB 181 EKRSTSRQFVGGPPPVKKTSTISGAVDQGNESIYAKTIVYVPPNDGPIEAVSTIETVP 240
OY 241 YWTRSRKKTGTLOPMNSDSTLNSROLEPRTETDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
DB 241 YWTRSRKKTGTLOPMNSDSTLNSROLEPRTETDSVGTPOSGNGMRLHDFVSKTVIKPESC 300
OY 301 VPCGRKIKFGKLSLKCRCDRVYSHPECRDRCPLPCIPILIGIPVKIGEGMLADVFVQSOTSP 360
DB 301 VPCGRKIKFGKLSLKCRCDRVYSHPECRDRCPLPCIPILIGIPVKIGEGMLADVFVQSOTSP 360
OY 361 MIPSTIVHGVNEIEORGLTETGLYRISGCDRTVYKELKEKFLKRYVPLLSKYDDIHAICS 420
DB 361 MIPSTIVHGVNEIEORGLTETGLYRISGCDRTVYKELKEKFLKRYVPLLSKYDDIHAICS 420
OY 421 LKDFLRLNKEPLTFRLNRAFMEEAETIDEDNSIAAMYQAVGELPQANRDTLAFIMIH 480
DB 421 LKDFLRLNKEPLTFRLNRAFMEEAETIDEDNSIAAMYQAVGELPQANRDTLAFIMIH 480
OY 481 QRYAOSPHTKMDVANLAKVGPPIVAHAVPNDPVTMSODIKRQKVERRLSLPLEYMS 540
DB 481 QRYAOSPHTKMDVANLAKVGPPIVAHAVPNDPVTMSODIKRQKVERRLSLPLEYMS 540
OY 541 QPMAMVEQENIDPLHYENSNASTPOTPRIKVSLGLPVTTPHQLLKTSSSSLSQRYMS 600
DB 541 QPMAMVEQENIDPLHYENSNASTPOTPRIKVSLGLPVTTPHQLLKTSSSSLSQRYMS 600
OY 601 TLTKMTPRSGSKSATNLGROGNFPA5PMLK 632
DB 601 TLTKMTPRSGSKSATNLGROGNFPA5PMLK 632

RESULT 3
AAU69422
ID AAU69422 standard; Protein; 570 AA.
XX
AC AAU69422;
XX
DT 30-JAN-2002 (first entry)
XX
DE Lung small cell carcinoma antigen #16.
XX
KM Human; cytostatic; antitumour; lung small cell cancer antigen;
XX
KW tumour; lung cancer.
XX
OS Homo sapiens.
XX
PN WO200177168-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001MO-US11859.
XX
PR 11-APR-2000; 2000US-196780P.
PR 21-JUN-2000; 2000US-213361P.
PR 01-SEP-2000; 2000US-229763P.
PR 05-SEP-2000; 2000US-230629P.

```

```

PR 14-SEP-2000; 2000US-232565P.
PR 19-DEC-2000; 2000US-257037P.
PR 08-JAN-2001; 2001US-260796P.
XX
XX (CORI-) CORIXA CORP.
XX
PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
XX
DR WPI: 2002-010986/01.
XX
DR N-PSDB; AAS61862.
XX
PS Claim 2; Page 282-284; 295pp; English.
XX
CC The invention relates to novel isolated lung small cell cancer antigen
CC polynucleotides (I) and polypeptides (II) used in a method of detecting
CC cancer in a patient. The method is optionally performed by
CC utilising oligonucleotides (III), where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAU69407-AAU69431 represent novel human lung small cell
CC cancer antigen amino acid sequences of the invention.
XX
SQ Sequence 570 AA;
XX
Query Match 90.0%; Score 2919; DB 23; Length 570;
Best Local Similarity 99.8%; Pred. No. 1.5e-236;
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 63 TERSALDVKLKHARNOVDVEIKRQRAEADCEKLEROIOLIREMLMCDTSGSIOLSEEQ 122
DB 1 TERSALDVKLKHARNOVDVEIKRQRAEADCEKLEROIOLIREMLMCDTSGSIOLSEEQ 60
OY 123 SALAFNLRGOPSSSSNAGNKRLLSTIDSGSILSDISFDKTESLDMDSSLVYKTFKKR 182
DB 61 SALAFNLRGOPSSSSNAGNKRLLSTIDSGSILSDISFDKTESLDMDSSLVYKTFKKR 120
OY 183 RSTSRQFVGGPPPVKKTSTISGAVDQGNESIYAKTIVYVPPNDGPIEAVSTIETVP 242
DB 121 RSTSRQFVGGPPPVKKTSTISGAVDQGNESIYAKTIVYVPPNDGPIEAVSTIETVP 180
OY 243 TRSRKKTGTLOPMNSDSTLNSROLEPRTETDSVGTPOSGNGMRLHDFVSKTVIKPESC 302
DB 181 TRSRKKTGTLOPMNSDSTLNSROLEPRTETDSVGTPOSGNGMRLHDFVSKTVIKPESC 240
OY 303 CGKRIKFGKLSLKCRCRCRVYSHPECRDRCPLPCIPILIGIPVKIGEGMLADVFVQSOTSP 362
DB 241 CGKRIKFGKLSLKCRCRCRVYSHPECRDRCPLPCIPILIGIPVKIGEGMLADVFVQSOTSP 300
OY 363 PSTIVHGVNEIEORGLTETGLYRISGCDRTVYKELKEKFLKRYVPLLSKYDDIHAICS 422
DB 301 PSTIVHGVNEIEORGLTETGLYRISGCDRTVYKELKEKFLKRYVPLLSKYDDIHAICS 360
OY 423 KDFLRLNKEPLTFRLNRAFMEEAETIDEDNSIAAMYQAVGELPQANRDTLAFIMIH 482
DB 361 KDFLRLNKEPLTFRLNRAFMEEAETIDEDNSIAAMYQAVGELPQANRDTLAFIMIH 420
OY 483 VAOSPHTKMDVANLAKVGPPIVAHAVPNDPVTMSODIKRQKVERRLSLPLEYMS 542
DB 421 VAOSPHTKMDVANLAKVGPPIVAHAVPNDPVTMSODIKRQKVERRLSLPLEYMS 480

```

QY 543 MAVEENIDPLHVIENSNAESTPOTPDIKVSLGPTTPEHQLKTPSSSSLSQVRSTL 602  
Db 481 MAVEENIDPLHVIENSNAESTPOTPDIKVSLGPTTPEHQLKTPSSSSLSQVRSTL 540  
QY 603 TKMTPRGSKSKSATNLGRGNFFASPMK 632  
Db 541 TKMTPRGSKSKSATNLGRGNFFASPMK 570

## RESULT 4

AAB94223  
ID AAB94223 standard; Protein: 255 AA.

AC AAB94223;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:14587.

KM Human; primer: detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saico K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PS Claim 8; SEQ ID 14587; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX Sequence 255 AA;

Query Match 36.4%; Score 1182; DB 22; Length 255;  
Best Local Similarity 99.2%; Pred. No. 6e-91;  
Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 397 KEKFLRVKTVPLSKVDHIAICSLKDFLRNKEPLLRLNRAFEAEITDEDSIA 456  
Db 20 KLFELRVKTVPLSKVDHIAICSLKDFLRNKEPLLRLNRAFEAEITDEDSIA 79  
QY 457 AMTQAVGELPOANDTLAFIMHLOVRAQSPHRTKMDVANAKVFGPTVAHAAPNPVT 516  
Db 80 AMTQAVGELPOANDTLAFIMHLOVRAQSPHRTKMDVANAKVFGPTVAHAAPNPVT 139  
QY 517 MSODIKRQPKVVERLSLPLEYWSQFMVQENIDPLHVIENSNAESTPOTPDIKVSLG 576  
Db 140 MLODIKRPKVERLSLPLEYWSQFMVQENIDPLHVIENSNAESTPOTPDIKVSLG 159  
QY 577 PVTTPHQLKTPSSSSLSQVRSTLTKMTPRGSKSKSATNLGRGNFFASPMK 632  
Db 200 PVTTPHQLKTPSSSSLSQVRSTLTKMTPRGSKSKSATNLGRGNFFASPMK 255

## RESULT 5

AAU32069  
ID AAU32069 standard; Protein: 256 AA.

AC AAU32069;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #2560.

XX Human; vaccination; gene therapy; nutritional supplement;

KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2001; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSED INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

PS Claim 20; Page 551; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising

CC the nucleic acids encoding the polypeptides and cells genetically

CC engineered to express them are also useful for producing the proteins.

CC The proteins are useful in genetic vaccination, testing and

CC therapy, and can be used as nutritional supplements. They may be used to

CC increase stem cell proliferation; to regulate haematopoiesis; and in

CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;

CC immune suppression and/or stimulation; as anti-inflammatory agents; and

CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 256 AA;

Query Match 30.0%; Score 971.5; DB 22; Length 256;

Best Local Similarity 84.0%; Pred. No. 3.1e-73;

Matches 205; Conservative 3; Mismatches 27; Indels 9; Gaps 4;

QY 357 QTSPIRISIVHCVNEIQRGLTEGLYRISGCDRTVKELKPLRKVTPLLSVVDIH 416  
 DB 1 QTSPIRISIVHCVNEIQRGLTEGLYRISGCDRTVKELKPLRKVTPLLSVVDIH 60  
 QY 417 AICSLKDFELRLKPLTEFLRLNRAFMFAETTDNDNSIAAMYQAVGELPOANRPTIAFL 476  
 DB 61 AICSLKDFELRLKPLTEFLRLNRAFMFAETTDNDNSIAAMYQAVGELPOANRPTIAFL 120  
 QY 477 MIHLQRAVQSPHTKMDVANLAKVFGPTIVAAHVNPDPVYTSQDIKR-QPKVERLLSLP 535  
 DB 121 MIHLQRAVQSPHTKMDVANLAKVFGPTIVAAHVNPDPVYTSQDIKR-QPKVERLLSLP 180  
 QY 536 L-EYVQPMVYEQE---NIDPLAVIENSNASTPOTPDIKV---SLGCVTPTPRHQLIK 587  
 DB 181 LMEYWFESFMDGMEQGTGFDPLHVIENSNAESTPOTPDIKAVPGGGLCVHFTAGBAEIQK 240  
 QY 588 TPSS 591  
 DB 241 GPPS 244

RESULT 6  
 ABB68056

ID ABB68056 standard; Protein; 625 AA.

AC ABB68056;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 30960.

KM Drosophila: developmental biology; cell signalling; insecticide;

OS pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL12159.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PR interactions -

XX Disclosure; SEQ ID NO 30960; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72022).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 625 AA;

Query Match 26.6%; Score 864; DB 22; Length 625;

Best Local Similarity 35.1%; Pred. No. 1.5e-63;

Matches 228; Conservative 110; Mismatches 225; Indels 86; Gaps 21;

QY 4 MNLVNRNLEQVLRVEILISEGN-EVQFVQLAKDEPDRKKQRTDHELGKXKDLMAE 62  
 DB 1 MALSLASEDDLRCKMOVLDTGTPPEEPLRFPLMEYOHEKACAGTAETARLQNELDSL 60  
 QY 63 TERSALDYKLRHARQVVEIKRRQAEADCEKLEROIQLIREMLCDTSGSIQSEBOK 122  
 DB 61 TKMGDLEGLFARHIDMEIKARQAEHERDAMESKIMAVADLLRHERN---LNNEFR 116  
 QY 123 SALAFLNQGPSSS-----NA--GNKRSLIDESSIISDSIFDKT-DESLDMSLYKT 174  
 DB 117 DKLAFLH-TPSSRRKSLINAVREDKSYGDINSTGLSLSTHSSEDFLD-----VKT 170  
 QY 175 FLKKRKRKRSTSRQFV-----DGPFGPKKTR--SIGSAVD----- 209  
 DB 171 SK-SWREHRPSSLPKQIPSPVGNKRRLSTGLNGMSGTTPPTGKRRSSVGIQVQHYVD 229  
 QY 210 --QGNESIVAKTTVTVPNDG-GPIEAVSTIETVPYWRSRKKTGTLQPMNSDSTLNSROL 266  
 DB 230 VQGAERCAATKVTIIPQDGGGVIRAEISTESLPIAINGERIGDGLSTPRRSVLEKATA 289  
 QY 267 EPRRTDSY-----GTP-QSNGMRHDFVSTVYKPESCVCGKRIKGLSKCR 317  
 DB 290 PELTPVNMAMPVVAESGTPLOHRPLRMHHTFSOKTFLRGDNCVOCORIRIRGAVGLRCR 349  
 QY 318 DCRVVSHPCCRDRCPICPFLIGTP-VKIGBMLADPVSQSPMIPISIVHCVNEIOR 376  
 DB 350 DCPVCHIDCKRLTLVSCVPO-TGTPPTKTMGYTDRAPSIAPMIPALIVHCVNEIAR 408  
 QY 377 GLTETGLYRISGCDRTVKELKPLRKVTPLLSVVDIHAIASLLKDFELRLKPLTEFL 436  
 DB 409 GLTEVGLYRLSSEREYKALKQFLRGKATPHLGMT-DIYVLCVCCVDFELRLKPLTEFL 467  
 QY 437 RLNRAPMFAAETTDNDNSIAAMYQAVGELPOANRPTIAFLMIHLQRAVQSPHTKMDVANL 496  
 DB 468 SQMKDFANAVQNPDPRTADMLVKSQKLPQANRPTIAFLMIHLQRAVQSPHTKMDVANL 527  
 QY 497 AKVFGPTIVAAHVNPDPVYTSQDIKRQPKVERLLSLPLEVYQPMVYEQEENIDPLHYI 556  
 DB 528 SLIRGPTIVGISTPPDDHAIYTEVFTQKQVKKALLELPVSWEDQIV-----IDPTR-- 580  
 QY 557 ENSNAESTPQT-----PDIKVSLGCVTPTP-----EHLQKTPSSSS 593  
 DB 581 -----TPATVYIKRVPSNKNKNDLLSLVATPEFGTIRKKRKFCTPPASA 622

RESULT 7

ABG14787

ID ABG14787 standard; Protein; 4318 AA.

XX ABG14787;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14778.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.





Best Local Similarity 46.9%; Pred. No. 8.1e-43;  
Matches 153; Conservative 35; Mismatches 68; Indels 70; Gaps 11;

```

OY 255 WNSDSLNSRQLEPRFETDSVGTPOSNGMRHLDPYKVIKPESCVPGCKRIKFEKLSL 314
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 369 WIKD--LVNRPRTIKTLENTGNTIODIGM--KDFMSKT---PKMATYTKIDKMDLIOL 422
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 315 KCRDCRVVSHPECRDRCPLPCIPITLIGTPVKIG-----EGMLADVQSOTSPMIPSIYVH 368
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 423 KSEFC-----TAKETIRVNRKOPTMEKIFATYSSD-----KGLISR 459
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 369 CVNEIEQ-----RGLTETGLY---RISGCDR--TYKELKEK-FLRY 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 460 IYNELKQIYKKKTNPNPKKWTNDMNRHFSKEDIYAAKKHMKKCSLPAIREMQIKTMYR 519
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 404 KTVPELL-----SKYDDIHATCSLKDPLRNKKEPLIFRLRAEAEITDEDN 453
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 520 HLPVRIAIIRKSGNNSKVDDIHATCSLKDPLRNKKEPLIFRLRAEAEITDEDN 579
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 454 SIAMVQAVGELPOANRDTLAFMLHLOVVAOSPHTKMDVANLAKVEGPTIYHAHPNDP 513
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 580 SISAMVQAVGELPOANRDTLVFLMHLQVVAOSPTTKMNVANLAEVFGSTIYAHAVNPE 639
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 514 PVTMSODIRKRPKVERLLSLPLEYV 539
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 640 PVTMLQDIKQCPKVERLPSLPLEYV 665
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

## RESULT 9

ABG19904  
ID ABG19904 standard; Protein; 1086 AA.

AC ABG19904;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19895.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR MPI: 2001-639362/73.

DR N-PSDB: AAS84091.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

XX Claim 20; SEQ ID No 50263; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1086 AA:

Query Match 19.1%; Score 618; DB 22; Length 1086;

Best Local Similarity 46.9%; Pred. No. 1.8e-42;

Matches 153; Conservative 35; Mismatches 68; Indels 70; Gaps 11;

```

OY 255 WNSDSLNSRQLEPRFETDSVGTPOSNGMRHLDPYKVIKPESCVPGCKRIKFEKLSL 314
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 790 WIKD--LVNRPRTIKTLENTGNTIODIGM--KDFMSKT---PKMATYTKIDKMDLIOL 843
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 315 KCRDCRVVSHPECRDRCPLPCIPITLIGTPVKIG-----EGMLADVQSOTSPMIPSIYVH 368
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 844 KSEFC-----TAKETIRVNRKOPTMEKIFATYSSD-----KGLISR 880
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 369 CVNEIEQ-----RGLTETGLY---RISGCDR--TYKELKEK-FLRY 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 881 IYNELKQIYKKKTNPNPKKWTNDMNRHFSKEDIYAAKKHMKKCSLPAIREMQIKTMYR 940
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 404 KTVPELL-----SKYDDIHATCSLKDPLRNKKEPLIFRLRAEAEITDEDN 453
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 941 HLPVRIAIIRKSGNNSKVDDIHATCSLKDPLRNKKEPLIFRLRAEAEITDEDN 1000
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 454 SIAMVQAVGELPOANRDTLAFMLHLOVVAOSPHTKMDVANLAKVEGPTIYHAHPNDP 513
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1001 SISAMVQAVGELPOANRDTLVFLMHLQVVAOSPTTKMNVANLAEVFGSTIYAHAVNPE 1060
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 514 PVTMSODIRKRPKVERLLSLPLEYV 539
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1061 PVTMLQDIKQCPKVERLPSLPLEYV 1086
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

## RESULT 10

ABG13958  
ID ABG13958 standard; Protein; 1139 AA.

AC ABG13958;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #13949.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.  
 DR N-PSDB: AAS78145.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID NO 44317; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at fip.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 XX Sequence 1139 AA;  
 SQ  
 Query Match 19.18; Score 618; DB 22; Length 1139;  
 Best Local Similarity 45.08; Pred. No. 1.9e-42;  
 Matches 158; Conservative 39; Mismatches 70; Indels 84; Gaps 14;  
 XX  
 QY 242 WTSRRRTGTLOPWNDSST-INSR-----OLEPTEEDSGTPOSGMRLH-----DF 289  
 DB 820 WLAICRKL-KLDPFLTVATINSRWIKDMLIRKT-----IKTLEBNIGITIDIGMKDF 874  
 QY 290 VSKTVIKPESCVPCGKRIKFKLSLKCRCRDVSVSHPCRDRCPLPIPTLIGPVKIG-- 347  
 DB 875 MSKT---PKAMATKTKDKMDLQKSFPC-----FAKETITRVNRQ 913  
 QY 348 -----EGMLADPVSQTSMIISIVHCHNEIQ-----RGLTEETLY- 384  
 DB 914 PTWEKIFATYSSD-----KGLISRIYNELKQIKKKTNNPIKKTNDNNRHSKEDIYA 968  
 QY 385 ---RISGCDR--TVKELKER-FLRVKTVPL-----SKVDIHAICSLIKDPLRN 428  
 DB 969 AKKHMKCSSLPRAIREQIKTYMRHYLTPYRIAIKKSGNKKVDDIHAICSLIKDPLRN 1028  
 QY 429 LKEPDLTFLRLNAPMEAEITDEDNSTAIAAYQAVGELPQANRDTLAFMLTLORVAPSPH 488  
 DB 1029 FKPEPLTFLRLNAPMEAEITDEDNSTAIAAYQAVGELPQANRDTLAFMLTLORVAPSPH 1088  
 QY 489 TKKDVANLAKVFGPTVAHAHVNPDPYNSODIKRQKVVYERLISPLEYV 539  
 DB 1089 TKMNVANLAEVFGSTVAHAHVNPDEPYMLDIIKQKVVYERLISPLEYV 1139  
 XX  
 RESULT 11  
 ID AAM41833 standard; Protein; 119 AA.  
 AC AAM41833;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 6764.  
 DE

XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 XX 26-JUL-2001.  
 XX  
 XX 26-DEC-2000; 2000WO-US34263.  
 XX  
 XX 21-JAN-2000; 2000US-0486725.  
 XX 25-APR-2000; 2000US-0552317.  
 XX 09-JUL-2000; 2000US-0598042.  
 XX 19-JUL-2000; 2000US-0620312.  
 XX 03-AUG-2000; 2000US-0653450.  
 XX 14-SEP-2000; 2000US-0662191.  
 XX 19-OCT-2000; 2000US-0693036.  
 XX 29-NOV-2000; 2000US-0727344.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu Z, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PSDB: AAI60989.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 XX Example 2; SEQ ID NO 6764; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 CC  
 XX Sequence 119 AA;  
 SQ  
 Query Match 17.94; Score 581; DB 22; Length 119;  
 Best Local Similarity 99.18; Pred. No. 6.5e-41;  
 Matches 115; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 MDTMMLNRMILFQOLVRRVETLSEGNVQFIQAKDFEDFRKKRQRTDHELKGYKDLAM 60  
 DB 4 MDTMMLNRMILFQOLVRRVETLSEGNVQFIQAKDFEDFRKKRQRTDHELKGYKDLAM 63  
 QY 61 AETERSALDYKLRHARNOVDVEIKRQRAEADCEKLERQIOLIREMLMCDTSGSII 116  
 DB 64 AETERSALDYKLRHARNOVDVEIKRQRAEADCEKLERQIOLIREMLMCDTSGSII 119  
 XX  
 RESULT 12  
 ABB59247

XX	ABBS59247 standard; Protein; 384 AA.
XX	
AC	ABBS59247;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 4533.
XX	
KW	Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.
XX	
OS	Drosophila melanogaster.
PN	
XX	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE ) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI: 2001-6556860/75.
XX	
DR	N-PSDB; ABL03350.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX	
PS	Disclosure; SEQ ID NO 4533; 21bp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072).
CC	
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 384 AA;

[illegible]

XX	RESULT 13
XX	AAM39659
ID	AAM39659 standard; Protein; 334 AA.
XX	
AC	AAM39659;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2804.
XX	
KW	Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropath; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
OS	Homo sapiens.
XX	
PN	MO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0486725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0596042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0723344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang Z, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
XX	
XX	N-PSDB; AAI58615.
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Example 4; SEQ ID NO 2804; 10078bp; English.
XX	
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: immune system suppression,
CC	Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
XX	Sequence 334 AA:
XX	
XX	Query Match 9.9%; Score 322; DB 22; Length 334;
XX	Best Local Similarity 32.0%; Pred. NO. 2,2e-18;
XX	Matches 90; Conservative 48; Mismatches 117; Indels 26; Gaps 8

Query Match	9.98;	Score 322;	DB 22;	Length 334;
Best Local Similarity	32.0%;	Pred. No. 2.2e-18;		
Matches	90;	Conservative 48;	Mismatches 117;	Indels 26;
				Gaps 8

```

Db      66 ATKENQIIPKYE-----KIHNEKVTFRGPHWCEYCA-NFMGLIAQGVKC 111
QY      317 RDCRVVSHPCRDRCPLPCIPITLIGTPVKIGEGMLADPVSOTSMISIVHCHNEIEQR 376
Db      112 ADGGLNVHKQCSKMPNDCKPDLKHVK-KVYSCDLTLVKAHTTKRPVVDKCIRESR 170
QY      377 GLTETGLYRISGCDRTYKELKEFLR--VKTVPLLSKVDDIHAICSLKDFLRNLKEPRL 434
Db      171 GLNSEGILYRVSQFSDLEDYKMAFDKRGKADISVNNYEDINITGALKLFLRDLPIPLI 230
QY      435 TFRLNRAFMEEAETTDENSIAMYOAVGELPOANRDTLAFMLTHLQVAVOSPHTK--M 491
Db      231 TYDAYPFFIESAKIMDPEDELETLEHAKLKLPPAHCELTLYLMAHLKRV--TLHEKENLM 288
QY      492 DVANLAKVFPPTIYAHAVPNPDPTVMSODIKRQKVVERLL 532
Db      289 NAENLGIVFGPTLMRS--PELDMAALNDIRYQRLVVELLI 327

```

# RESULT 14

AAW41445 ID AAW41445 standard; Protein: 352 AA.

AAW41445:

22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 6376.

Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

Homo sapiens.

WO20015312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

WPI: 2001-442253/47.

N-PSDB: AAI60601.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

Example 2: SEQ ID NO 6376; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAW38642-AAW42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide

of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 352 AA;

Query Match 9.9%; Score 322; DB 22; Length 352;

Best Local Similarity 32.08; Pred. No. 2,3e-18;

Matches 90; Conservative 48; Mismatches 117; Indels 26; Gaps 8;

QY 259 STLNSROLEPPTETDSVGTQDSNGMRLHDFVSKTVIKPSCVPCGRIRFGKLS--LKC 316

Db 84 ATKENQIIPKYE-----KIHNEKVTFRGPHWCEYCA-NFMGLIAQGVKC 129

QY 317 RDCRVVSHPCRDRCPLPCIPITLIGTPVKIGEGMLADPVSOTSMISIVHCHNEIEQR 376

Db 130 ADGGLNVHKQCSKMPNDCKPDLKHVK-KVYSCDLTLVKAHTTKRPVVDKCIRESR 188

QY 377 GLTETGLYRISGCDRTYKELKEFLR--VKTVPLLSKVDDIHAICSLKDFLRNLKEPRL 434

Db 189 GLNSEGILYRVSQFSDLEDYKMAFDKRGKADISVNNYEDINITGALKLFLRDLPIPLI 248

QY 435 TFRLNRAFMEEAETTDENSIAMYOAVGELPOANRDTLAFMLTHLQVAVOSPHTK--M 491

Db 249 TYDAYPFFIESAKIMDPEDELETLEHAKLKLPPAHCELTLYLMAHLKRV--TLHEKENLM 306

QY 492 DVANLAKVFPPTIYAHAVPNPDPTVMSODIKRQKVVERLL 532

Db 307 NAENLGIVFGPTLMRS--PELDMAALNDIRYQRLVVELLI 345

## RESULT 15

AAW75995 ID AAW75995 standard; Protein: 1261 AA.

AAW75995:

18-NOV-1998 (first entry)

GTPase activating protein (GAP), PARG.

PARG; GTPase activating protein; GAP; PTP1L polypeptide; phosphatase;

Rho family signal transduction; cancer cell; proliferation; GAP domain;

Ras-like GTPase; Rho; Rac; Cdc42; Pdz domain; signal transduction;

intracellular protein tyrosine phosphatase; mast cell secretion;

pharmacological agent; modulator; diagnosis.

Homo sapiens.

Key Location/Qualifiers

FT Domain 193..509

FT Domain 613..652

FT Domain 658..898

FT Domain 9705-0805363.

FT Domain 9705-0805363.

FT Domain 9705-0805363.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Aspenstrom P, Franzen P, Gopez LJ, Heldin C, Hellman U,  
PI Saras J;

DR WPI; 1998-467559/40  
DR N-PSDB; AAV55846.

DR N-PSDB; AAV55846.

PT Polypeptide, PARC, with GTPase activating activity - useful e.g.  
PT therapeutically to reduce Rho family signal transduction so reduce  
PT cancer cell proliferation, modulate mast cell secretion

PS Claim 18; Pages 41-48; 93pp; English

CC This represents a human GTPase activating protein (GAP), PAR6. The  
CC specification states that the unique PAR6 nucleic acid molecule  
CC specifically excludes known sequences shown in AA055847 to AA055850.  
CC  
CC Expression vectors comprising the PAR6 nucleic acid can be used to  
CC transform host cells for the recombinant production of the protein. The  
CC PAR6 polypeptides or complexes comprising PP1L polypeptide having  
CC phosphatase activity (AA075939) and PAR6 polypeptide can be administered  
CC to mammalian cells to reduce Rho family signal transduction. The PAR6  
CC polypeptides can be administered to cancer cells to reduce proliferation  
CC (which is increased by Rho family signal transduction). Alternatively,  
CC dominant-negative variants of the polypeptide (e.g. with a GAP domain  
CC inactivated) can be used to increase Rho family signal transduction.  
CC PAR6 was determined to have GAP activity in vitro, which was strongest  
CC Rho but also detected on Ras-like GTPases, Rac and Cdc42, and bound  
CC specifically to the fourth PDZ domain (PDZ4) of known intracellular  
CC protein tyrosine phosphatase PP1L, involved in signal transduction.  
CC PAR6 polypeptides and polynucleotides are useful to produce modulators of  
CC PAR6 GTPase activity useful to reduce Rho family signal transduction or  
CC modulate mast cell secretion. The polypeptides are also useful for  
CC identifying pharmacological agents useful for disease diagnosis and  
CC treatment.

**SQ Sequence 1261 AA;**

Query Match	Score	DB	Length
9.38;	301;	DB 19;	Length 1261,

best local similarity 22.4%; Pred. NO. 1.1e-157  
Matches 153; Conservative 120; Mismatches 245

Matches 153; Conservative 120; Mismatches 249; Indels 160; Gaps 33;

```

QY      5 RNLFEQYVRVELSEGNVQFIOIAKPFEDRPRKMOQTHDE-IGYKODLIMKETSBSA 67
Db      372 RLREELALOKVE---EADBLKYV-CVTYNEERANDVEYTKREIIAQLRLTVFOCDLTJKA
QY      68 LDVYKLAHANNQYDVEIKRRORAEADCEKLEQRIOLIREMLMDTSGSIOLSEOK---SA 124
Db      428 VYVNFLEFMQHLOAASIALDRLOSLCSAKLYDPOGEYSFEVKAATNS-----TEEEKYDGNV
QY      125 LAFNLNGQSSSNAGNKLRLSTI---DEGSLTSDISFDKDESIDMDSSLYKPFKLKRE 181
Db      483 NKHLNLSQPSGFGPANSLEDVYRLDPDSNNKLTIEORCSNSADIT---GPFSRLSWTFMGFS
QY      182 KRRTSRQFVDPRPPVYKTRTSIGSVAQNGENSIYAKTTVYVPDNGRIEAVSTIEVPY 241
Db      540 DSEST-----GGSSRSRLSDS-----EST-----SPED-----562
QY      242 WTRSRKRTGTLOPWNDSSTLNSRQLEPTEP--DSVGTPOS--NGGMRLHDFSVYK 296
Db      563 FHKKLPRTPTSSGTMSADLDREPRPSPSETGPNLSIFKTKTLMSKAALTHKF--RLTRS 620
QY      297 PESCVPGCRIFGFKLSKCRDRCRVYSEPCRDQPLPC-----LP---TLIGPVYKIGEG 349
Db      621 PTKRCBCBEIVF---QGVCECECLVCHRKCLENLVYICGHQKILPGKIHILFG-----670
QY      350 MLADF--VSQTSR-MIPSIYVCHVEIQRGLTETGLYRISGCDRTVKELEKELRYKTV 406
Db      671 --AETFLVAKKBPDPIDPIFIKICASEIENRALCLQGIYRVCG-----NKKTE 716
QY      407 PLLSKVD-----DI-----HATGSLADFLRNKKEPLTRRLNRAEAAE---ITDE 451
Db      717 KICLALENGMHLIDSEFSSHDICVYALVIRQPEPIFLRKPEIDILAKETIOHNE 776

```

```

OY      452 DNS-----IAAMTQAVELEQOARDJLAEIMJLQLOVA -OSPTHT 490
Db      777 QETKKSLSDECKMPNMCIEINRLILKSKDOLKQLPASFNSLHFLYHLKRYVDHAENK 836
OY      491 MDVANLAKYFEGPTIVAAHVPNDP--PYTMSQ--DIKROPKYVERLLSLPLEYMSQFM-- 543
Db      837 MNSKNLGVLEFGPLIR---PRQPTAPITISSLAEYSNQARLVEFLIT-----YSQKIFDG 888
OY      544 -MVEQENIDPLAVIENS---NAFSPPQPPDIKVSILSGVYTPPEHQLLKTPSSSSLSQVR 599
Db      889 SLQPOVMSISGVDDGCEPPKDLLSPBEERDIERSMKLSFFSKEDIHTESESKIFERAT 948
OY      600 STLTKNTPRGSKSKSATNLGR 621
Db      949 S-----FEESERQONALGK 962

```

Search completed: July 3, 2003, 09:54:08  
Job time : 76 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:52:57 ; Search time 43 seconds

(without alignments) 1412.952 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMLNLRNLEFQLVRRVE.....SKSATNLGRGQNFASPMLK 632

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3243	100.0	632	2	D59430
2	547.5	16.9	681	2	T23454
3	477	14.7	383	2	B48122
4	477	14.7	384	2	A48122
5	348	10.7	1980	2	S54307
6	334	10.3	2022	2	A59256
7	329	10.1	334	2	S29128
8	322	9.9	299	2	S08242
9	320	9.9	299	2	A43953
10	308	9.5	299	2	T39954
11	301	9.3	1261	2	E59430
12	294	9.1	295	2	S29956
13	294	9.1	443	2	A53764
14	293.5	9.1	466	2	B53764
15	291	9.0	970	2	D59435
16	290.5	9.0	814	2	F59430
17	286.5	8.8	1944	2	A59438
18	285.5	8.8	1165	2	D59433
19	282	8.7	859	2	A49307
20	281	8.7	822	2	B47485
21	281	8.7	822	2	A47485
22	280.5	8.6	733	2	S44876
23	275.5	8.5	1271	1	TVHDBR
24	270.5	8.3	1397	2	T46354
25	270.5	8.3	2548	2	E59435
26	269.5	8.3	735	2	A59434
27	267.5	8.2	2626	2	T31099
28	260.5	8.0	802	2	H59434
29	258	8.0	903	2	T00705

## ALIGNMENTS

30	253.5	7.8	1846	2	T33079	hypothetical prote
31	247.5	7.6	655	2	A59430	hypothetical prote
32	247	7.6	863	2	T27958	hypothetical prote
33	245	7.6	1445	2	A59437	KIAA1204 protein l
34	240	7.4	969	2	T38478	RhoGAP/PLM domain
35	238	7.3	974	2	E59434	Rho GTPase activat
36	232	7.2	666	2	S29349	hypothetical prote
37	231.5	7.1	647	2	A57467	RalBp1 - rat
38	231	7.1	818	2	A59433	KIAA0672 protein l
39	229.5	7.1	574	2	T29005	hypothetical prote
40	229.5	7.1	655	2	F59435	Rala-binding prote
41	229	7.1	634	2	T27959	hypothetical prote
42	228.5	7.0	837	2	T19825	hypothetical prote
43	222.5	6.9	892	2	T40040	GTPase-activator p
44	221.5	6.8	316	2	T46471	hypothetical prote
45	217	6.7	512	2	E59437	F02569_2 protein l

## RESULT 1

D59430  
Rac GTPase activating protein 1 [imported] - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 19-Jul-2002  
C:Accession: D59430  
R:Kawashima, T.; Kitamura, T.; Nosaka, T.; Hirose, K.  
Submitted to GenBank, December 1999  
A:Description: Homo sapiens Rac GTPase activating protein 1 (RACGAP1), mRNA.  
A:Reference number: D59430  
A:Accession: D59430  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-632 <KAW>  
A:Cross-references: GB:NP\_037409; PID:G7019433; PIDN:NP\_037409.1

## Query Match

100.0%; Score 3243; DB 2; Length 632;

Best Local Similarity 100.0%; Pred. No. 1.9e-188;

Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDTMLNLRNLEFQLVRRVEILSEGENEVOFIQADFEFRKKMORTHELKGYKDLMMK	60
DB	1	MDTMLNLRNLEFQLVRRVEILSEGENEVOFIQADFEFRKKMORTHELKGYKDLMMK	60
QY	61	AETERSALDVKLKHARNQVDVEIKRQRAEADCEKLEKQIOLIREMLCMTSGSIQISEE	120
DB	61	AETERSALDVKLKHARNQVDVEIKRQRAEADCEKLEKQIOLIREMLCMTSGSIQISEE	120
QY	121	OKSALAFINRGQSSNMGKRLSTIDSGSLISIDKDESDMDSLSVTKFKLKR	180
DB	121	OKSALAFINRGQSSNMGKRLSTIDSGSLISIDKDESDMDSLSVTKFKLKR	180
QY	181	EKRRTSRQFVDGPPGPKTRTSIGSAVDGNESTVARTTVVPPNDGPGIEAVSTIETVP	240
DB	181	EKRRTSRQFVDGPPGPKTRTSIGSAVDGNESTVARTTVVPPNDGPGIEAVSTIETVP	240
QY	241	YWTSSRRKTGTLOPWNSTLNSRQLEPRTETDSVGTQOSNGMRLHDFSVTKIPESC	300
DB	241	YWTSSRRKTGTLOPWNSTLNSRQLEPRTETDSVGTQOSNGMRLHDFSVTKIPESC	300
QY	301	VPCGRIKFGKLSKCRQCRVVSHEPCDRCPILPILIGPVVIGSGMLADVFQSPSP	360
DB	301	VPCGRIKFGKLSKCRQCRVVSHEPCDRCPILPILIGPVVIGSGMLADVFQSPSP	360
QY	361	MIPSIIVHCVNEIEQGLTEGLYRISGCDRTVKELEKFKLVKTVPLLSKYVDIHAICS	420
DB	361	MIPSIIVHCVNEIEQGLTEGLYRISGCDRTVKELEKFKLVKTVPLLSKYVDIHAICS	420
QY	421	LIKDFLRNLKEPLTLFRLNRAFMFAETTEDDNGSIANYQAVGELPQANRDTLAFMLIHL	480
DB	421	LIKDFLRNLKEPLTLFRLNRAFMFAETTEDDNGSIANYQAVGELPQANRDTLAFMLIHL	480

QY 481 QRAVQSPHTKMDVANKAFVGPPIVAHAANVPDPVTMSODIKRQPKVERLLSLPLEYWS 540  
 DB 481 QRAVQSPHTKMDVANKAFVGPPIVAHAANVPDPVTMSODIKRQPKVERLLSLPLEYWS 540  
 QY 541 QFMMVQENIDPLHVIENSAFSTPOTPOLKVSLLGPVTPPEHOLKTPSSSSLSQVR 600  
 DB 541 QFMMVQENIDPLHVIENSAFSTPOTPOLKVSLLGPVTPPEHOLKTPSSSSLSQVR 600  
 QY 601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632  
 DB 601 TLTKNTPRFGSKSKSATNLGROGNFPASPMLK 632

## RESULT 2

T23454  
 hypothetical protein K08E3.6 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C/Accession: T23454  
 R/McMurray, A.  
 submitted to the EMBL Data Library, November 1996  
 A/Reference number: Z19743  
 A/Accession: T23454  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-681 <MIL>  
 A/Cross-references: EMBL:Z81568; PIDN:CA04593.1; GSPDB:GN00021; CESP:K08E3.6  
 A/Experimental source: clone K08E3  
 C/Genetics:  
 A:Gene: CESP:K08E3.6  
 A:Map position: 3  
 A/Introns: 36/1; 73/2; 237/3; 361/3; 612/3

Query Match 16.9%; Score 547.5; DB 2; Length 681;  
 Best Local Similarity 26.0%; Pred. No. 4,8e-27;  
 Matches 101; Conservative 109; Mismatches 284; Indels 121; Gaps 22;

QY 7 NVNLFEOALVRRVELISEGNEVOFIOLANDEFEDFRKK-----ORTHELKYYKDL 59  
 DB 14 NSHHIIMILINSQPOFDIKDIFMLIDEIERLRLKMKDESKRKLNDMREAEALA 73  
 QY 60 KATERSALDYKAKHARNOVDKIKRORAEADCEKLEIOLIRELMCDTSGSLOLSE 119  
 DB 74 KARKKRLAMDIDVDQKHLRALMEENKALKLDLVYETREKQKAKMKGIFNS--LTK 131  
 QY 120 EOKSALAFNRGQSSNAGNKRILSTIDSGSILSDISFDKTESLDMW---SSLVYTFK 176  
 DB 132 EDRDQKFLHE---PLVRYTSKRVQ--QRHPIHMETQDDEDESEVDYDETGSFEVYH 186  
 QY 177 LKK-REKRSTST-----ROFVDDGPPG---PVKK 200  
 DB 187 LRRGRFVRSSAAGNAVGGKRSASAHATTAANSKRSRGVTATIDEEPNGGPPKR 246  
 QY 201 TYSIGAVNOGNSIYAKTIVT-----VPNDGSP-----IEAVSTIETVP 240  
 DB 247 CRDDGSTPRQEMTTTTTTTTTTHNSRAONODPPRVSILHROLTRSLSGSIPSCDQTP 306  
 QY 241 YWTRSRRTGTLQPMNSDSTLNSROLEPRTETDVGTPGOSNGMRLLHDFVSKTVIPESC 300  
 DB 307 GQTTNNIGLMSAIIITKSTLDIRLTKRGTPANTNGTTR-DIMRPHPTFEACIKAMKRC 365  
 QY 301 VPGKRIRKFGKLSLKCRCGRVYSHPCRCRCPJCP-----TLIGTPVKIGEGM 350  
 DB 366 DMCATATLKL-A-TSMKRCRCHQVYHRSCKMLHPCIPRPKTMTPKSAALGARAGGEPR 424  
 QY 351 LADFSQTSMPPIPSIVYHCYNEIEQGLTETGLYRISGCDRYVKEKEFLRVKYVPLS 410  
 DB 425 LQDFCSAPRMPPIPAAYIHCYVALEARGLTQEGIRYRPGQVRYVNVLLDE-LRSKTYPNVG 483  
 QY 411 KYDDIHAICSLKDFRNKKEPLTFRLNRAFWEAELI--TDEDNSTIAAYQAVGELOPA 468  
 DB 484 -LHDEVITDYLKRLFKLDKLDPLIPRTSRQELIVANLSTDPDNGRLNLRVCELOPA 542

QY 469 NRDTLAFMLTHLQV-AQSPHTKMDVANKAFVGPPIVAHAANVPDPVTMSQ-----D 520  
 DB 543 NRDTLAFMLTHLQV-AQSPHTKMDVANKAFVGPPIVAHAANVPDPVTMSQ-----D 520  
 QY 521 IKRQPKVERLLSLPLEYWSQFMMVQENIDPLHVIENSAFSTPOTPD-----IKYSLG 576  
 DB 597 ATDCHRMATLAFEEVDYVWQRFJGTA-----VSNASQIETARHODFALCDRSLIG 649  
 QY 577 PATTPEHOLKTPSSSSLSQVRSTLTNTPRFGS 611  
 DB 650 PATT-----SPATPLANSANATRAARGHLLGS 677

## RESULT 3

B48122  
 GTPase-activating protein Rac homolog, splice form clone pcl.7d - fruit fly (Drosophila)  
 C/Species: Drosophila melanogaster  
 C/Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Feb-1997  
 C/Accession: B48122  
 R/Aguel, M.; Roder, L.; Voia, C.; Griffin-Shea, R.  
 Mol. Cell. Biol. 12, 5111-5122, 1992  
 A/Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal  
 A/Reference number: A48122; M01D:93024458; PMID:1406685  
 A/Accession: B48122  
 A/Status: preliminary  
 A/Molecule type: nucleic acid  
 A/Residues: 1-383 <AGN>  
 A/Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIP:115663)  
 C/Genetics:  
 A:Gene: FlyBase:rn  
 A/Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;  
 Best Local Similarity 40.7%; Pred. No. 6,5e-23;  
 Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

QY 279 QSNNG-MRLHDEVSKT-VIKPESCPCGRIKFKGLSLKCRDGVVSHPCRDRCPLCT 336  
 DB 78 QSHSGILREHNFIRKISYYVNGCVHCRKIRPAMSLRCRACPLCHIGCCQLTYVNCI 137  
 QY 337 P-TLIGTPVYIGGMLADPVSQSPMIPSIYVHCYNEIEQGLTETGLYRISGCDRTYKE 395  
 DB 138 PGPQIGK---RGCSLDYAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPV 193  
 QY 396 LKEKFLRVKTVPLSKYDDHAIKSLKDFRLMKJELTFRLNRAFWEAELTDEDNSTI 455  
 DB 194 LRRKLKRGKSTPHLGN-KDHTTLCOCVQKDLQVHPLIPYHRRDFEETRHEDLAVE 252  
 QY 456 AAMYQAVGELPQANRDTLAFMLTHLQV-AQSPHTKMDVANKAFVGPPIVAHAANVPDPV 515  
 DB 253 MAYVLAVLELHQHNRDTLAFMLTHLQV-AQSPHTKMDVANKAFVGPPIVAHAANVPDPV 515  
 QY 516 TMSODIKRQPKVERLLSLPLEYWSQFMMV 545  
 DB 307 LTLENVATWQRIKLVLLMPAGFWSQFLEV 336

## RESULT 4

A48122  
 GTPase-activating protein Rac homolog, splice form clone pcl.7 - fruit fly (Drosophila)  
 C/Species: Drosophila melanogaster  
 C/Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Feb-1997  
 C/Accession: A48122  
 R/Aguel, M.; Roder, L.; Voia, C.; Griffin-Shea, R.  
 Mol. Cell. Biol. 12, 5111-5122, 1992  
 A/Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal  
 A/Reference number: A48122; M01D:93024458; PMID:1406685  
 A/Accession: A48122  
 A/Status: preliminary  
 A/Molecule type: nucleic acid  
 A/Residues: 1-384 <AGN>  
 A/Note: sequence extracted from NCBI backbone (NCBIN:115660, NCBIP:115661)  
 C/Genetics:





```

QY 71 ----KLHNAQVDEIKRROBARADEKLEROLILREMLMCPOTSSIOJSEOGKSLA 126
Db 1459 PSGOHRRAAGE-----KRTKEGGCKKKNNKVI-----GKTIYSEKWRSEVF 1502Z
QY 127 FLANGOPSSSNAGNKKRLSTIDE-SGSILSDISPEKTDSELDMSLWTFELKREKRRS 185
Db 1503 ----RQITMNA-----ELKYLDEFILNKINDROSKTP-----IESLEIATEKFRS 1545Z
QY 186 TSROFVDPGPGPVKRTTSGISGAVDOGNEISIAKTTVTVPNDGPIEAVSTIET-VPYWTR 244
Db 1546 NIKTMYSVPNGKI-----HWGKXDLMEYQIIVVSMILATERGOKDNTLVLFNFQSLDEPTR 1601Z
QY 245 SRRTGTGLPQPNSSSTILNSRQDERLETEDSVGTQOSNGMMLHDFVSKTYIAKPSYPCG 304
Db 1602 GYTT-----NDEPQVQSAQKKRKKRQERAVOEHNG---HWFASYOVSIPQSDOCL 1650Z
QY 305 KRIKFGKSLKRCRCRVRVSHPECRCRCPRLPIPL-IGRP-VKIGE-GMLADPVSQTS 360
Db 1651 SYTLMOKALLCSYCKMKCHKKCYHKIKIOSHCSYTYGKKGEGEARGHGVGVDSLTSDKA 1710Z
QY 361 MIPSIYHCVNEIEORGLTETGLYRISGCDTVKLEKFLRYKTVPLSKVDD-IHAI 418
Db 1711 SVPIVLEKLEHVEHMGHLYTEGLYRKSGAANRTRELQ--ALQTDPAVYLVENFPIHAI 1767Z
QY 419 CSLKLDPLRNKKEPILFRLNRARMEAEITDENSTIAAMTQAVGELPOANRPLATLMT 478
Db 1768 TGVLCQWMLRELPETLMTFAOYGDLELRAVELPEKEOEOLAAIYAVLEHNPBEAHHNSLERLIF 1827Z
QY 479 HLOHVA-OSPHTKMDVANIARVFEPTVIAHVPPMDPVTMSODIKROKPYVERLSLEPLE 537
Db 1828 HLVAVALLEDVNRKSPGALAIIFAPCLL-ROPDMSDPLTSSKDVKILITTCVEMILIKQMR 1886Z
QY 538 TWQSFPMWVEDENIDPLHVIEN-----SNAFSPPOQPD-----IKVSLGQVT 579
Db 1887 KYK-----VKMEIISQLEAEIESIAFRRLSILRONANKSPKTEPAPGAGRLLTTSRVSP-- 1940Z
QY 580 TPEHOALKTPSSSLSQVRST-----LTKNPFR 608
Db 1941 SPSTRNIALGWSRAALRTGTGPBARBARALRRRPPR 1980

```

RESULT 7  
S29128  
N:Chimerin - rat  
N:Alternate names: GTPase-activating protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 25-Feb-1994 #sequence\_revision 19-Oct-1995 #text\_change 08-Oct-1999  
C:Accession: S29128; S25152  
R:Llm, H.H.; Michael, G.J.; Smith, P.; Llm, L.; Hall, C.  
Biochem. J. 287, 415-422, 1992  
A:Title: Developmental regulation and neuronal expression of the mRNA of rat n-chimerin  
A:Reference number: S29128; MUID:93074974; PMID:1445199  
A:Accession: S29128  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-334 <LIM>  
A:Cross-references: EMBL:X67250; NID:g55939; PIDN:CAA4672.1; PID:g55940  
R:Llm, H.; Michael, G.J.; Smith, P.; Llm, L.; Hall, C.  
submitted to the EMBL Data Library, July 1992  
A:Description: Rat n-chimerin a p21rac GAP: cDNA sequence developmental regulation and r  
A:Reference number: S25152  
A:Accession: S25152  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-334 <LIM>  
A:Cross-references: EMBL:X67250; NID:g55939; PIDN:CAA4672.1; PID:g55940  
C:Superfamily: protein kinase C zinc-binding repeat homology <KK2>  
;81-130/Domain: protein kinase C zinc-binding repeat homology <KK2>

```

OY 255 STNSROLEBRETDSVGRPOSGMBRLHDEFSKIVIPESCVQGRKIKFGKTS--LKC 316
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 AILKEBOLPKYE-----KANEVNIHFRPHMCEYCA-NEMMGLIAOGVC 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 317 RDCRVVSHPECRDRCPILPILIGTPVKGIGMLADFSOTSPMIBSVHCNEIEQR 376
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 ADGGLANVHKQCSKMYNPNDCKPDLKNHK-KVYSCDLTTLVKAHITKRPVWDMCIRESESR 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 377 GLNETGLYRSGCDRIVKELKEKPLR--KVYPLILSKYDDIHAJCSLLKODLRNKLPEPL 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 GLNSEGLIKRSGSDLIEDYKMAFDRDGEKADISVNMEDINIIITGALKILFYRLPLPLI 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 435 TFLRNRAFMBAEITDEDSIAAMYQAVGELPOANRDTLAEMLILORVAPSPTK--M 491
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 TYDAYPFEISAKIVPDEQLETHLEALSLSPALCEFLRYLMAHLKRV--TLHEKENLM 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 492 DVANIKVREPTIVAHAVNPDPVPMVSQOIKRQPFVVERLL 532
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 SAENLIGVIEGPTLMRS--PELDPMALANDIRIORIVLELLI 327

```

RESULT 8  
S08242  
N:Chimerin - human  
N:Alternate names: GTPase-activating protein  
C:Species: Homo sapiens (man)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 05-Nov-1999  
C:Accession: S08242  
J:Hall, C.; Monfries, C.; Smith, P.; Lhm, H.H.; Kozma, R.; Ahmed, S.; Vanniasingham,  
J. Mol. Biol. 215, 11-16, 1990  
A:Title: Novel human brain cDNA encoding a 34,000 M(r) protein n-chimerin, related to  
A:Reference number: S08242; MUID:90135942; PMID:225965  
A:Accession: S08242  
A:Molecule type: mRNA  
A:Residues: 1-299 <HA1>  
A:Cross-references: EMBL.X51408; NID:G35012; PIDN:CA035769.1; PID:G35013  
C:Superfamily: protein kinase C zinc-binding repeat homology  
F:46-95/Domain: protein kinase C zinc-binding repeat homology <K2>

Query Match	9.9%	Score 322	DB 25	Length 299
Best Local Similarity	32.0%	Pred. No. 3.3e-13		
Matches	90	Conservative	48	Mismatches 117
			Indels	26
			Gaps	8
QY	259	STLSNRQLPEPTREIDSVGTPOSGNGMRLLHDFYSKVIYIPESCVPGGRKRIKEKLS--LKC	316	
DB	31	ATLKENEIDIPYE-----KIHNEVFHFRGPWHECYA-NFMWLLIAAGVNC	76	
QY	317	RDCRVSHPECRDRNCPICPIPTLLIGTPVKIGEGMADVSQSPMIPSIYVHCVEIOR	376	
DB	77	ADCGLNVAHKQSCWMPNDNCKPDLKIVK-KVYSCDLTTLVKAHTTRPVMVDMCIREISER	135	
QY	377	GLTEGGLYRIGCDRTVELEKEKFLR--VKIYPLLSKYDDHIALTSCLKDFLRNKEPL	434	
DB	136	GLNSEGLRVSGFSDLIEDVKMAFPRDEGEKADI SVNME DINITIGALKLYFRDLPIL	195	
QY	435	TFRLRAFMEEALETIDEDNSIAAQAQVAGELPOANRDILAFILMIHQRYAOSPHTK--M	491	
DB	196	TYDAVPKRTESAKIMDPDEQLETLHEALKLLPPACHCETLRYLMAHLKRV--TLHEKENIM	253	
QY	492	DVANLAKYFGPTIYAHAVPNPDPTVMSODIKQRPVYERLL	532	
DB	254	NAENIGIVFGPIIMRS--PELDAMALNDIRQRLVLELLI	292	

RESULT 9  
A43953  
N-Chimerin - common canary  
M:Alternate names: protein kinase C homolog [misidentification]; song control circuit  
C:Species: *Serinus canaria* (common canary)  
C:Date: 26-Aug-1999 #sequence\_revision 26-Aug-1999 #text\_change 04-Feb-2000  
C:Accession: A43953  
R:George, J.M.; Clayton, D.F.  
Brain Res. Mol. Brain Res. 12, 323-329, 1992



```

Db      671  -AEFLVAKKEDGJFIFLIKACSEIENFALCIQGIYRVCG-----NKIKTE 716
QY      407  PLISKVD-----DI-----HAICSLKDFLRNKEKELFLFRNRAEAE-----ITDE 451
Db      717  KILCALENGMLHVLDISESSHIDCVLKLRLQPEPFLFRKEFDLAKIOTAHNEE 776
QY      452  DNS-----IAAMYQAVGELPQANRDTLAFIMTLHQRVA-QSPHTK 490
Db      777  QETKKNSELEDKKMPNMCIEINRILKSKDLRLQAPANFMSIHFLYHLAKRYVHAENK 836
QY      491  MDVANLAKVFGPTVAHAVNPDP--PYTMSQ--DIKQPKYVERLSLPLEYMSQFM--- 543
Db      837  MNSKRNIGVIGPSPILIR---PRPQAPITITISLAEYSQARLVEFLIT-----YSQKIFDG 888
QY      544  -WYEOENIDPLHYIENS---NAFSTPQTPDIKYSILGAPVTTPHQHLKTPSSSLSQRVR 599
Db      889  SLDPQDVMCSIGVYDQCCFPKPLISPERRDIERSMKSLFFSKEDITHTSSESKIFERAT 948
QY      600  STLTAKTPRFGSKSKSATNLGR 621
Db      949  S-----FESESRKQNALGK 962

```

## RESULT 12

```

beta-chimerin - rat
A:Accession: A45485
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <LEU>
A:Cross-references: EMBL:X69489; NID:957526; PDB:CAA49244.1; PID:957527
A:Experimental source: testis
A:Note: sequence extracted from NCBI backbone (NCBI:125731)
R:Leung, T.; How, B.E.; Manser, E.; Lim, L.
Submitted to the EMBL Data Library, November 1992
A:Description: Germ cell beta-chimerin, a new GTPase-activating protein for p21rac, is
A:Reference number: S31398
A:Accession: A45485
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <LEU>
A:Cross-references: EMBL:X69462; NID:955816; PID:955817
C:Superfamily: protein kinase C zinc-binding repeat homology
F:42-91/Domain: protein kinase C zinc-binding repeat homology <K22>

```

```

Query Match      9.18; Score 294; DB 2; Length 295;
Best Local Similarity 29.38; Pred. No. 1.9e-11;
Matches 81; Conservative 52; Mismatches 111; Indels 32; Gaps 8;

```

```

QY      285  RLHDFVSKYTYKPESCVPCGKRIFKGLS--LKCRCDRVYSHPECRCRCLPCLPTLIGT 342
Db      40  KTNHFKVHTIRGPHWCYCA-NFMWGLIAGVRSDDGLVNHKQCSKHVPNDCCPDLKRI 98
QY      343  PVKIGEGMLADFYQSOTSPMIPSIYHCVNEIEQRLGTETGLYRISGCDRTVKELKEFLR 402
Db      99  K-KVYCCDILTLYKANTQRPVYVDCIRLEIARGLKSEGLYRVSGETHEIEDVKMAFDR 157
QY      403  VKTVPLLSKYD-----DIHAICSLKDFLRNKEKELFLFRNRAEAEITDEDSNI 455
Db      158  DG-----EKADISANIYPDINIITGALKLTFRDLPIPIIYDYTKFIEAKISNADERL 212
QY      456  AAMYQAVGELPQANRDTLAFIMTLHQRVAOSPHRK-MDVANLAKVREPTVAHAVNPDP 514
Db      213  EAVHEVLMPLPRAHYETLRIMHLKRYTNKERNKDNLMANLNLGIVEGPTLM--RPPDST 270

```

```

QY      515  VTMSODIKRQPKYVERLSLPLEYMSQFMWEOENI 550
Db      271  LTTLHDKRYOKLIVQ-----ILLNEDV 293

```

## RESULT 13

```

beta2-chimerin, cerebellar - rat (fragment)
A:Accession: B53764
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-443 <LEU>
A:Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F:34-114/Domain: SH2 homology <SH2>
F:190-239/Domain: protein kinase C zinc-binding repeat homology <K22>

```

```

Query Match      9.18; Score 294; DB 2; Length 443;
Best Local Similarity 29.38; Pred. No. 3.4e-11;
Matches 81; Conservative 52; Mismatches 111; Indels 32; Gaps 8;

```

```

QY      285  RLHDFVSKYTYKPESCVPCGKRIFKGLS--LKCRCDRVYSHPECRCRCLPCLPTLIGT 342
Db      188  KTNHFKVHTIRGPHWCYCA-NFMWGLIAGVRSDDGLVNHKQCSKHVPNDCCPDLKRI 246
QY      343  PVKIGEGMLADFYQSOTSPMIPSIYHCVNEIEQRLGTETGLYRISGCDRTVKELKEFLR 402
Db      247  K-KVYCCDILTLYKANTQRPVYVDCIRLEIARGLKSEGLYRVSGETHEIEDVKMAFDR 305
QY      403  VKTVPLLSKYD-----DIHAICSLKDFLRNKEKELFLFRNRAEAEITDEDSNI 455
Db      306  DG-----EKADISANIYPDINIITGALKLTFRDLPIPIIYDYTKFIEAKISNADERL 360
QY      456  AAMYQAVGELPQANRDTLAFIMTLHQRVAOSPHRK-MDVANLAKVREPTVAHAVNPDP 514
Db      361  EAVHEVLMPLPRAHYETLRIMHLKRYTNKERNKDNLMANLNLGIVEGPTLM--RPPDST 418
QY      515  VTMSODIKRQPKYVERLSLPLEYMSQFMWEOENI 550
Db      419  LTTLHDKRYOKLIVQ-----ILLNEDV 441

```

## RESULT 14

```

beta2-chimerin, cerebellar - human
A:Accession: A53764
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-466 <LEU>
A:Cross-references: GB:I29126; NID:9457229; PDB:AAA19191.1; PID:9457220
C:Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F:57-137/Domain: SH2 homology <SH2>
F:213-262/Domain: protein kinase C zinc-binding repeat homology <K22>

```

```

Query Match      9.18; Score 293.5; DB 2; Length 466;
Best Local Similarity 25.78; Pred. No. 4e-11;
Matches 98; Conservative 66; Mismatches 155; Indels 63; Gaps 12;

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:44:41 ; Search time 25 Seconds

(without alignments)  
1048.522 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MOTMLNLRNLFEEQLVRRVE.....SKSATNLGRGNFASPMLK 632

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	477	14.7	384	1 RN_DROME	P40809 drosophila
2	348	10.7	1980	1 MY9B_RAT	Q63508 rattus norv
3	339.5	10.5	2114	1 MY9B_MOUSE	Q99Y06 mus musculu
4	329	10.1	334	1 CHIN_RAT	P30337 rattus norv
5	328	10.1	334	1 CHIN_MOUSE	Q91457 mus musculu
6	323.5	10.0	2158	1 MY9B_MOUSE	Q13459 homo sapien
7	322	9.9	459	1 CHIN_HUMAN	P15882 homo sapien
8	294	9.1	295	1 CHIO_RAT	Q03070 rattus norv
9	293.5	9.1	468	1 CHIO_HUMAN	Q02757 homo sapien
10	290.5	9.0	814	1 OPHL_HUMAN	Q09131 mus musculu
11	282	8.7	859	1 ABR_HUMAN	Q12979 homo sapien
12	280.5	8.6	747	1 YN54_CAEEL	P34588 caenorhabd
13	275.5	8.5	1271	1 BCR_HUMAN	P11374 homo sapien
14	260.5	8.0	802	1 OPHL_HUMAN	Q06090 homo sapien
15	251.5	7.8	986	1 RHG6_MOUSE	Q54834 mus musculu
16	244	7.5	802	1 OPHL_MOUSE	Q99313 mus musculu
17	238	7.3	974	1 RHG6_HUMAN	Q43182 homo sapien
18	232	7.2	666	1 RGDL_YEAST	P38339 saccharomyc
19	228.5	7.0	837	1 YLBS_CAEEL	P46541 caenorhabd
20	206.5	6.4	1275	1 YA09_SCHPO	Q10164 schizosacch
21	202.5	6.2	946	1 RHG4_HUMAN	P98171 homo sapien
22	197.5	6.1	638	1 Y053_HUMAN	P43231 homo sapien
23	192	5.9	718	1 RHG8_HUMAN	Q9nsg0 homo sapien
24	191	5.9	425	1 RHG8_MOUSE	Q9c8p4 mus musculu
25	188.5	5.8	439	1 RHG1_HUMAN	Q07960 homo sapien
26	187	5.8	2167	1 BEM2_YEAST	P39960 saccharomyc
27	180	5.6	622	1 Y3B1_HUMAN	Q9Y313 homo sapien
28	171	5.3	650	1 Y411_HUMAN	Q43295 homo sapien
29	169.5	5.2	1007	1 RGAL_YEAST	P39983 saccharomyc
30	169	5.2	601	1 Y3B1_MOUSE	P55194 mus musculu
31	160.5	4.9	1017	1 LRGI_YEAST	P35588 saccharomyc
32	159.5	4.9	1128	1 BEM3_YEAST	P32873 saccharomyc
33	142	4.4	728	1 P85B_HUMAN	Q00459 homo sapien

## ALIGNMENTS

RESULT 1	ID	RN_DROME	STANDARD:	PRT:	384 AA.
AC	P40809	Q9V191.			
DT	01-FEB-1995	(Rel. 31, Created)			
DT	01-FEB-1995	(Rel. 31, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	GTPase activating protein rRacGAP.				
GN	RNRACGAP OR CG2595.				
OC	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;				
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	STRAIN-Oregon-R; TISSUE=Pupae;				
RC	MEDLINE=93024458; PubMed=1406685;				
RA	Agnel M., Roeder L., Vola C., Griffin-Shea R.;				
RT	"A Drosophila rotund transcript expressed during spermatogenesis and				
RT	imaginal disc morphogenesis encodes a protein which is similar to				
RT	human Rac GTPase-activating (racGAP) proteins."				
RL	Mol. Cell. Biol. 12:5111-5122(1992).				
RL	[2]				
RP	SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.				
RC	MEDLINE=96194886; PubMed=8654933;				
RA	Heemann C.D., Bergeret E., Guichard A., Griffin-Shea R.;				
RT	"Alternative splicing of the Drosophila melanogaster rotundracGAP				
RT	gene."				
RL	Gene 168:135-141(1996).				
RL	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	STRAIN=Berkeley;				
RC	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherter S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.C., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,				
RA	wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Adair J.F., Agbayani A., An H.-C., Andrews-Plankkoc C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bohlav S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,				
RA	Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				

34	137.5	4.2	3093	1	POLG_BSTVL	065730 b genome po
35	137	4.2	724	1	P85B_BOVIN	P23726 bos taurus
36	135	4.2	1395	1	CUT1_MOUSE	P53564 mus musculu
37	135	4.2	3678	1	DMD_MOUSE	P11531 mus musculu
38	134	4.1	988	1	PCK1_SCHPO	P36582 schizosacch
39	132.5	4.1	724	1	P85A_BOVIN	P23727 bos taurus
40	132	4.1	705	1	YNP9_CAEEL	P34562 caenorhabd
41	130.5	4.0	1080	1	HDA4_CHICK	P83038 gallus gall
42	130.5	4.0	3685	1	DMD_HUMAN	P11532 homo sapien
43	129	4.0	1505	1	CUT1_HUMAN	P39680 homo sapien
44	128.5	4.0	878	1	KPCO_HUMAN	Q9b216 homo sapien
45	128.5	4.0	1616	1	P200_MYCCE	Q49429 mycoplasma

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Mlishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Port V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -1- FUNCTION: INVOLVED IN THE MORPHOGENESIS OF THE ADULT APPENDAGES.  
 CC GTPASE-ACTIVATING PROTEIN FOR P21-RAC. PROMOTES THE EXCHANGE OF  
 CC RAC-BOUND GDP BY GTP.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: IN PUPAE, EXPRESSED IN IMAGINAL DISKS AND ONLY  
 CC IN THE MALE GONAD. IN ADULTS, ONLY FOUND IN THE TESTES, IN REGIONS  
 CC FILLED WITH PRIMARY SPERMATOCYTES.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMATOGENESIS, IN PRIMARY  
 CC SPERMATOCYTES, AND IMAGINAL DISK MORPHOGENESIS.  
 CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG  
 CC BINDING DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-23 IS THE INITIATOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M9702; AAA16480.2; -;  
 DR EMBL: U22539; AAC47075.1; -;  
 DR EMBL: U22539; AAC47076.1; -;  
 DR EMBL: AE003672; AAF54034.1; -;  
 DR PIR: A48122; A48122.  
 DR FlyBase: FBgn0045843; rRacGAP.  
 DR InterPro: IPR002219; DAG\_PE\_bind.  
 DR InterPro: IPR000198; RhoGAP.  
 DR Pfam: PF00130; DAG\_PE\_bind.1.  
 DR Pfam: PF00620; RhoGAP.1.  
 DR SMART: SM00109; C1; 1.  
 DR SMART: SM00324; RhoGAP.1.  
 DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM.1; 1.  
 DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM.2; 1.  
 CC Guanine-nucleotide releasing factor; Spermatogenesis; Testis;  
 CC Alternative splicing; Zinc; Phorbol-ester binding.  
 CC AM DOMAIN 87 136 PHORBOL-ESTER AND DAG BINDING.  
 CC FT DOMAIN 188 202 ARG/LYS-RICH.  
 CC FT VASPLIC 378 384 NLSSTHL -> RPYMSL (IN ISOFORM 2).  
 CC SEQUENCE 384 AA; 44217 MW; D24959B833B7A2AB CRC64;  
 CC -----  
 QY Query Match 14.7%; Score 477; DB 1; Length 384;  
 CC Best Local Similarity 40.7%; Pred. No. 2e-23;  
 CC Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;  
 QY 279 OSNG-MRLHDFVSKT-VIKESGCPGCKRIKFKLSKCDKCRVSHPEKDRCPKCI 336  
 DB 78 QSHSGGLREHNFKIKSYIVNGNCVHCCKRIREFMASTKRCACPRLCHIGCCROLTVNCI 137  
 QY 337 P-TLIGPVKIGEGMLADFEVSQTSMPISYIVHCVNEIEGRLTRETGLRISGCDRYKE 395  
 DB 138 PQDQIGTK----RCGISDYAPRAVAPWALYVHCYTELEAGLQDQEGIRYSSSTRECKR 193  
 QY 396 LKEFLKRVKTVPLLSKVDIHAICSLKDLRLNKEPLTLFRLRNRAPEAAETDDEDSI 455

DB 194 LRRKLIRKSPHIGN-KDTHTLCQVDFEFLQVLPIPIYHRDEFEARHEDRLAVE 252  
 QY 456 AAMQAVELQANDTLAFLMIHQRVASPHKMDVNAKVGPIVAHVPNDPV 515  
 DB 253 MAVVIAVELQAHDDTLAYLMLIMOKIAESPAAVMTVNNLAVIFAPLFG-----DLD 306  
 QY 516 TMSODIKRPKVERLSTLPLEYMSOFMNV 545  
 DB 307 LTLLEVYQWQVIAVLLIMPRGPFMSQFLEV 336  
 CC -----  
 CC RESULT 2  
 CC MY9B\_RAT  
 CC ID MY9B\_RAT STANDARD; PRT; 1980 AA.  
 CC AC 063358;  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Myosin IXb (Unconventional myosin-9b).  
 CC GN MYO9B OR MYR5.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC OX NCBI\_TaxID=10116;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN-Sprague-Dawley; TISSUE-Brain stem, and spinal cord;  
 CC RX MEDLINE=9518874; PubMed=7882973;  
 CC RA Reinhard J., Scheel A.A., Diekmann D., Hall A., Ruppert C.,  
 CC Baehler M.;  
 CC RT "A novel type of myosin implicated in signalling by rho family  
 CC GTPases.";  
 CC RL EMO J. 14:697-704(1995).  
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE  
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.  
 CC MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS  
 CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP  
 CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.  
 CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.  
 CC -1- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLocalizes WITH  
 CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS  
 CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE  
 CC PERINUCLEAR REGION (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, LUNG, THYMUS, BRAIN,  
 CC LIVER, AND SPLEEN.  
 CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG  
 CC BINDING DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 IQ DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X77609; CA54700.1; -;  
 DR HSSP: P08799; IIMD.  
 DR InterPro: IPR002219; DAG\_PE\_bind.  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR000159; RA\_domain.  
 DR InterPro: IPR000198; RhoGAP.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head.2.  
 DR Pfam: PF00130; DAG\_PE\_bind.1.  
 DR Pfam: PF00612; IQ.4.  
 DR Pfam: PF00620; RhoGAP.1.  
 DR Pfam: PF00788; RA.1.  
 DR PRINTS: PR00193; MYOSINHEAVY.



ProDom: PD000355; myosin\_head; 2.  
 DR SMART: SM00109; CL: 1.  
 DR SMART: SM00015; IO: 3.  
 DR SMART: SM00242; MYSC: 1.  
 DR SMART: SM00314; RA: 1.  
 DR SMART: SM00324; RHOGAP: 1.  
 DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1: 1.  
 DR PROSITE: PS00081; DAG\_PE\_BIND\_DOM\_2: 1.  
 DR PROSITE: PS00096; IO: 3.  
 DR PROSITE: Repeat; ATP-binding; Calmodulin-binding; Actin-binding;  
 KW Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;  
 KW Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;  
 KW Zinc.  
 FT DOMAIN 1 940 HEAD OR MOTOR DOMAIN.  
 FT DOMAIN 941 1045 NECK OR REGULATORY DOMAIN.  
 FT DOMAIN 1046 1980 TAIL.  
 FT DOMAIN 1841 1861 COILED COIL (POTENTIAL).  
 FT DOMAIN 1918 1948 COILED COIL (POTENTIAL).  
 FT DOMAIN 845 856 ACTIN-BINDING.  
 FT DOMAIN 958 978 IO 1.  
 FT DOMAIN 981 1001 IO 2.  
 FT DOMAIN 1002 1024 IO 3.  
 FT DOMAIN 1025 1054 IO 4.  
 FT DOMAIN 1593 1641 PHORBOL-ESTER AND DAG BINDING.  
 FT DOMAIN 1673 1822 RHO-GAP.  
 FT NP\_BIND 239 ATP (POTENTIAL).  
 FT NP\_BIND 246 ATP (POTENTIAL).  
 SQ SEQUENCE 1980 AA; 225035 MW; D79FEC4D0FAEC05 CRC64;

Query Match 10.7%; Score 348; DB 1; Length 1980;  
 Best Local Similarity 22.1%; Pred. No. 3.3e-14;  
 Matches 160; Conservative 105; Mismatches 262; Indels 196; Gaps 25;

82 EKRRORADDECKERLOILR--EMLMCDTSGSLQSEKSAFLINRGP---SSS 136  
 1272 QIORIO--HPDERLATAVEINRGKKLASAMLSQSLDSKPTAGALPTERRISFS 1329  
 137 NAGNKRSLTIDSGSLDISFDK-----TDESIDMDSLVYTKR 176  
 1330 TSDVSKLSPVKTSTVDGDLSAKKRPAKHKSSDPASGPDAGLPISQSDSKSAFRLEL 1389  
 177 LKREKRRS-----TSRQFYDGP-----PGPVKTRSI 204  
 1390 HNAKDKPFLSEVEETEGSGGQAADAPARKTLDPVSSQDHRHTTEKPLKCKKNRKY 1449  
 205 G--SAVDGNESTIVAKT-----VTVPND---GGPIEA-----VS 234  
 1450 GQITVSEKRESVFKITIANELKFLDELFLANKVNDLSOKPTIESLFTAEERFRSNIK 1509  
 235 TLETVP-----YTTSRRKGT 251  
 1510 TMYSVNGKIHGYKIDMENYQIVSNLAERGENKTNLVNFQSLDLEPTFSYKNT-D 1568  
 252 LQPMNSDTLNSROLEPRTETDSVGTPOSGCRLHDEVSKYVAKESCPGCKRIKFGK 311  
 1569 FEVVKGAOKKKRKOERAVO-----EHNG---HVFASYVNIPQSEOCISYIWMLD 1617  
 312 LSLKRCDCVNSHPECRDRCPLPIPT-----LIGTPYKISGMADVVSQSPMIPSI 365  
 1618 KALLCSVCMKCHKKCVHKIQSYCYTGRRKSELGAEPRGF--GVCVDSTLSKASVPYV 1675  
 366 VVHCVNEIQRGLTETGLYRISGCDRTVRELKEKFRKVTVPPLTSFYVD--IHAISCLK 423  
 1676 LKLELHEVMHGILYETGLYRKSGAANRTRELQ--ALQTDPAIVGLDEPPIHAIIGVYK 1732  
 424 DELRLNKEPLTLFRNRAFMEAEITDEDNSIAAMYQAVGELPQANRDTLAFIMILQRY 483  
 1733 QMLRELPEPLMTFAQYGDLELRAVELPEKOEOLAAIYAVLDHLPEANHTSLERLIFLVKY 1792  
 484 A-QSPHTKMDVANLAVPGPTIVAHAVPNDPVTMKODIKRQKVYVERLLSPLXWSQF 542  
 1793 ALLEDVNRMSPGALITFAFCLL-KCPDMSDPLTSMKVDLKITTCVEMLIKQEMRRYK-- 1849  
 543 MMEQENIDPLHYIEN-----SNA-----FSTPQRPDIKVSLLGPTVTPEHOH 585

Db 1850 --VKMEINHLAEASIAFRRLSLRQNAWPLKLGESSPYEGVRTKSPRTPVYQDLEL 1907  
 QY 586 LKTPSSSS-----LSQYRS-----TLTKTTPRGSKSKSATNLGROGNFASP 629  
 Db 1908 GALPEEAAGDEDEKEIKEMERIOSIKEREKIDITYRLPEIDPGSDENIDSETASTES 1967  
 QY 630 MKK 632  
 Db 1968 LLE 1970

RESULT 3  
 MY9B\_MOUSE  
 ID MY9B\_MOUSE STANDARD; PRT; 2114 AA.  
 AC Q9QY06; Q9QY07; Q9QY08; Q9QY09;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin IXb (Unconventional myosin-9b).  
 GN MYO9B OR MYR5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=20047919; PubMed=10580159;  
 RA Grewal P.K., Jones A.-M., Macconochie M., Lemmers R.J.F., Frants R.R.,  
 RA Hewitt J.E.;  
 RT "Cloning of the murine unconventional myosin gene Myo9b and  
 RT identification of alternative splicing."  
 RT Gene 240:389-398(1999).  
 RL  
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH APPASE  
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.  
 CC MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS  
 CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP  
 CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.  
 CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.  
 CC -1- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH  
 CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS  
 CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE  
 CC PERINUCLEAR REGION (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/Q AND 3/C ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN, HEART, MUSCLE AND  
 CC INNER EAR.  
 CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG  
 CC BINDING DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 IQ DOMAINS.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL: AF143687; AAF00122.1; -  
 EMBL: AF143685; AAF00120.1; -  
 DR EMBL: AF143686; AAF00121.1; ALT\_FRAME.  
 DR EMBL: AF143683; AAF00118.1; -  
 DR HSSP: P08799; 1MND.  
 DR MGD: MG1:106624; MYO9B.  
 DR InterPro: IPR002219; DAG\_PE-bind.  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR000159; RA\_domain.  
 DR InterPro: IPR000198; RHOGAP.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head; 2.





RT "Cloning of the murine unconventional myosin gene Myo9b and  
RT identification of alternative splicing."  
RT Gene 240:389-398(1999).  
RN [3]  
RP SEQUENCE OF 1940-2158 FROM N.A. (LONG ISOFORM).  
RX MEDLINE:98158729; PubMed:9490538;  
RA Post P.L., Bokoch G.M., Mooseker M.S.;  
RT "Human myosin-Ixh is a mechanotchemically active motor and a GAP for  
RT rho."  
RT J. Cell Sci. 111:941-950(1998).  
RN [4]  
RP SEQUENCE OF 1828-2023 FROM N.A. (SHORT ISOFORM).  
RC TISSUE-Placenta;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Ishihashi T., Fujimori K.,  
RA Tanai H., Kimata M., Matanabe M., Hiraoka S., Ishii S., Kawai Y.,  
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,  
RA Masuho Y., Kanehori K.;  
RT "NDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE  
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.  
CC MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS  
CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP  
CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.  
CC ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.  
CC -1- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLocalizes WITH  
CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS  
CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE  
CC PRINCIPAL REGION.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN PERIPHERAL BLOOD  
CC LEUCOCYTES AND AT LOWER LEVELS, IN THYMUS, SPLEEN, TESTIS,  
CC PROSTATE, OVARY, BRAIN, SMALL INTESTINE, AND LUNG.  
CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHOSBO-ESTER AND DAG  
CC BINDING DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- CAUTION: THE C-TERMINAL SEQUENCE FROM AA 1917 ONWARDS FROM REF. 1  
CC WAS PROBABLY A CHIMERA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U42391; AAC50402.1; ALT\_SEQ.  
DR EMBL: AF143684; AAF00119.1; -  
DR EMBL: AF020267; AAC26597.1; -  
DR EMBL: AK002201; BAA92132.1; ALT\_INIT.  
DR HSSP: P08799; IMND.  
DR Genew: HGNC:7609; MYO9B.  
DR MIM: 602129; -  
DR InterPro: IPR002219; DAG\_PE-bind.  
DR InterPro: IPR000048; IQ\_region.  
DR InterPro: IPR000159; RA\_domain.  
DR InterPro: IPR000198; Rhogap.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00063; myosin\_head\_2.  
DR Pfam: PF00130; DAG\_PE-bind\_1.  
DR Pfam: PF00612; IQ\_4.  
DR Pfam: PF00620; Rhogap\_1.  
DR Pfam: PF00788; RA\_1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR ProDom: PD000355; myosin\_head\_2.  
DR SMART: SM00109; C1; 1.  
DR SMART: SM0015; IQ; 3.  
DR SMART: SM00242; Mysc; 1.

DR SMART: SM00314; RA; 1.  
DR SMART: SM00324; Rhogap; 1.  
DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
DR PROSITE: PS00081; DAG\_PE\_BIND\_DOM\_2; 1.  
DR PROSITE: PS50096; IQ; 3.  
RW Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;  
KW Cytoskeleton; Coiled coil; GTPase activation; Phospho-ester binding;  
KZ Zinc; Alternative splicing.  
FT DOMAIN 1 939  
FT 940 1044 HEAD OR MOTOR DOMAIN.  
FT 1045 2158 NECK OR REGULATOR DOMAIN.  
FT 1046 1071 TAIL.  
FT 1880 1901 COILED COIL (POTENTIAL).  
FT 1959 1989 COILED COIL (POTENTIAL).  
FT 844 855 COILED COIL (POTENTIAL).  
FT 957 977 ACTIN-BINDING.  
FT 979 1000 IQ 1.  
FT 1001 1023 IQ 2.  
FT 1024 1053 IQ 3.  
FT 1633 1681 PHOSBO-ESTER AND DAG BINDING.  
FT 1713 1862 RHO-GAP.  
FT NP\_BIND 239 246  
FT VASPLIC 2022 2023 PP -> QY (IN SHORT ISOFORM).  
FT VASPLIC 2024 2158 MISSING (IN SHORT ISOFORM).  
FT CONFLICT 1937 1939 QVP -> KT (IN REF. 4).  
FT CONFLICT 1947 1947 L -> P (IN REF. 4).  
FT CONFLICT 2040 2045 TVAAP -> PMPPLH (IN REF. 3).  
FT CONFLICT 2049 2049 P -> L (IN REF. 3).  
FT CONFLICT 2067 2067 P -> S (IN REF. 3).  
FT CONFLICT 2157 2158 NG -> MAESHS (IN REF. 3).  
SQ SEQUENCE 2158 AA; 243556 MW; 4978F1D70F56D28 CRC64;

Query Match 10.0%; Score 323.5; DB 1; Length 2158;  
Best local similarity 22.3%; Pred. No. 1,3e-12;

Matches 149; Conservative 117; Mismatches 264; Indels 119; Gaps 27;

22 LSEGENVO-----FIQLAKDFEDFRKKWQ---RTDHELGYKYLDMRAETERSALDV 70.  
Db 1404 LSPSQVDSKSTFKRLFLHKTKD-----KRTSLGAELENAVSGHVLATYTKKGLA 1458  
71 ----KIKHARQVVEIKRRRAADCEKLERQQLIREMLMCTSGSIQSEOKSALA 126  
Db 1459 PSQOQHRRAGE-----KRTKEPGGKGRKNRWKI-----GKTVSEKWRSSVF 1502  
QY 127 FLNQGSSSNAGKRLSTIDE-SGSLTSDISPKDTESLDWDSLVKTFKKRRKRS 185  
Db 1503 ----RQITNAN-----ELKYDEFLNKINDRSKTP-----IESLFIENKFRS 1545  
QY 186 TSROFVDPGPEPVAKTRNSIGSAVDQNESTIVAKTTVPNDGPIEAVSTIET-VPYWR 244  
Db 1546 NIKTMYGVPNKIKI-----HVGKIDLMENYQIVASMLATERGOKDNLVNLFOSLIDEFTR 1601  
QY 245 SRRTGTLQLPWNSSTLNSRLERETEDSGTFQNSGKRLHDFVSKYIKKPCSCVPG 304  
Db 1602 GYTK-----NDEPVRKQSAQKRRKORAVOEHNG-----HVEASVQVSIPOSCQCL 1650  
QY 305 KRIKFGSLKRCRCRVVSHPECRDPLCIPLT-IGTP-VKIG-GMLADPVQSOTSP 360  
Db 1651 SYILMDKALLCSCKAKTKCHKYKIKQSICSTYTGKKGEGAPGHFGCVDSLTSDKA 1710  
QY 361 MIPSIIVHCVNEIEBQRLTETGLYRISGCDRTVYELKEKFLRYKVTPLSKVDN-THAI 418  
Db 1711 SVPIVLEKLEHVEHMGLEGLYKRSKANRTRRELQ---ALQTPAAVKLENFPIHAI 1767  
QY 419 CSLKDLRLNKEFLFLRLNRAFMEAEITDENSTIAAMYQAVGELPQANRPLALMI 478  
Db 1768 TGVLKQWLRELPELMTFAQYGFELRAVELPEKEQOJLAATYAVLEHPEANHNSLERLIF 1827  
QY 479 HLOEVA-QSPHTKMDVANLAKVGEPTVAANAVERPDVNTSODIKRQPKVYERLSLPLE 537  
Db 1828 HLYKVALLEDVNRKSPALAIITRAPCL-RCPPDSPLTSMKQDLAKTTTCEVEMLIKQMR 1886  
QY 538 YWSQFMNVEQENIDPLHVIEN-----SNA-----FSTPQTPDIKVSILGPVT 580



Db 356 TYADYPRFISAKIMDEOLETLHEALKLPAPHCETLRIMAHMKRV--TLHEKENIM 413

Oy 492 DVANLAKVFPTIVAHAVPNPDPTMSODIKROKPYVERL 532

Db 414 NAENLGIVFGPTLMRS--PELDAMALNDIRYORLVELLI 452

RESULT 8

CHIO\_RAT STANDARD: PRT: 295 AA.

ID CHIO\_RAT

AC 003070;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-chimerin (Beta-chimerin).

GN CHN2 OR BCH.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Testis;

RX MEDLINE=93179371; PubMed=8440677;

RT Leung T., How B.E., Manser E., Lim L.;

RT "Germ cell beta-chimerin, a new GTPase-activating protein for

RT p21rac, is specifically expressed during the acrosomal assembly stage

RT in rat testis.";

RL J. Biol. Chem. 268:3813-3816(1993).

CC -1- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC.

CC -1- SUBCELLULAR LOCATION: Membrane-associated (potential).

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BETA-1 (SHOWN HERE) AND BETA-2;

CC AAE PRODUCED BY ALTERNATIVE SPLICING. THE CEREBELLAR BETA-2

CC FORM CONTAINS A SH2 DOMAIN.

CC -1- TISSUE SPECIFICITY: FOUND IN CEREBELLUM AND TESTIS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN LATE STAGE

CC SPERMATOCTES. IN THE CEREBELLUM, EMERGENCE OF BETA-2 ISOFORM

CC COINCIDES WITH GRANULE CELLS MATURATION AND EXHIBITS POSTNATAL

CC DEVELOPMENTAL INCREASES. EXPRESSION IS SPECIFICALLY REDUCED IN

CC WEAVER MUTANT.

CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG

CC BINDING DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC

EMBL: X69489; CAA49244.1; -

EMBL: L07494; AAA40809.1; -

DR HSPB, P28867; IPTQ.

DR InterPro: IPR002219; DAG\_PE-bind.

DR InterPro: IPR000198; RhoGAP.

DR Pfam: PF00130; DAG\_PE-bind.1.

DR Pfam: PF00620; RhoGAP.1.

DR PRINTS: PR00008; DAGPEDOMAIN.

DR SMART: SM00109; C1.1.

DR SMART: SM00324; RhoGAP.1.

DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM.1; 1.

DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM.2; 1.

KW GTPase activation; Phorbol-ester binding; Zinc; Membrane; SH2 domain;

KW Alternative splicing.

FT DOMAIN 42 91 PHORBOL-ESTER AND DAG BINDING.

FT DOMAIN 118 262 RHO-GAP.

SO SEQUENCE 295 AA; 33837 MW; D7692D957B4816BD CRC64;

Query Match 9.1%; Score 294; DB 1; Length 295;

Best Local Similarity 29.3%; Pred. No. 6/7e-12;

Matches 81; Conservative 52; Mismatches 111; Indels 32; Gaps 8;

Oy 285 RLHDFVSTVTKPESCVPCGKRIFKGLS--LKRCDCRVVSHPCRCRCPPLCIPITLIGT 342

Db 40 KTHNFKXHTFGPFWCEYCA-NFMWGLIAGVRCSDGLNHNKCSNHPVNDQPDLRKI 98

Oy 343 PVKIGEGMLADFVSQTSMPISIVVHCVNEIORGLETETGLYRISGDRVYKELKEFLR 402

Db 99 K-KVYCCDLFTLVKAHNTQRMVVDICRIETEARGLKSEGLRYVSGFTEHIEDVKMAFDR 157

Oy 403 VKTVPPLSKVD-----DTHATCSLKDFLRNKEPFLPRLNARMEAAETIDENSI 455

Db 158 DG-----EKADISANIPVDINITGALKLFRDLPITPIYDYTKFTEAKISNAERL 212

Oy 456 AAMTQAVGELPOANRDTLAFIMHLOQVAPSPHTR-MDVANLAKVFPTIVAHAVPNPD 514

Db 213 EAVHEVLMPLPRAVETLRIMHILKVTYMKENKNNMNAENLGIVFGPTLM--RPPDST 270

Oy 515 VTMSODIKROKPYVERLLSLPLEYWSQPMWVEQENI 550

Db 271 LTLTHDMRYOKLIVQ-----ILLENDV 293

RESULT 9

CHIO\_HUMAN STANDARD: PRT: 468 AA.

ID CHIO\_HUMAN

AC P52757;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-chimerin (Beta-chimerin).

GN CHN2 OR BCH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Cerebellum;

RX MEDLINE=94230370; PubMed=8175705;

RT Leung T., How B.-E., Manser E., Lim L.;

RT "Cerebellar beta 2-chimerin, a GTPase-activating protein for p21

RT ras-related rac is specifically expressed in granule cells and has a

RT unique N-terminal SH2 domain.";

RL J. Biol. Chem. 269:12888-12892(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;

RX MEDLINE=95393937; PubMed=7614486;

RT Yuan S., Miller D.W., Barnett G.H., Hahn J.F., Williams B.R.G.;

RT "Identification and characterization of human beta 2-chimerin:

RT association with malignant transformation in astrocytoma.";

RL Cancer Res. 55:3456-3461(1995).

RN [3]

RP SEQUENCE OF 18-192 FROM N.A.

RA Strong C., Graves T., Yoakum M., Hawkins M.;

RA submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC. INSUFFICIENT

CC EXPRESSION OF BETA-2 CHIMERIN IS EXPECTED TO LEAD TO HIGHER RAC

CC ACTIVITY AND COULD THEREFORE PLAY A ROLE IN THE PROGRESSION FROM

CC LOW-GRADE TO HIGH-GRADE TUMORS.

CC -1- SUBCELLULAR LOCATION: Membrane-associated (potential).

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BETA-1 AND BETA-2 (SHOWN HERE);

CC ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE BRAIN AND PANCREAS. ALSO

CC EXPRESSED IN THE HEART, PLACENTA, AND WEAKLY IN THE KIDNEY AND

CC LIVER. EXPRESSION IS MUCH REDUCED IN THE MALIGNANT GLIOMAS.

CC COMPARED TO NORMAL BRAIN OR LOW-GRADE ASTROCYTOMAS.

CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG

CC BINDING DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L29126; AAA19191.1; -;  
 DR EMBL: U07223; AAA16836.1; -;  
 DR EMBL: U28926; AAA86528.1; -;  
 DR EMBL: AC004417; AAC06177.1; -;  
 DR HSSP: P28867; IPTO.  
 DR Genew: HGNC:1944; CHN2.  
 DR MIM: 602857; -;  
 DR InterPro: IPR002219; DAG\_PE\_bind.  
 DR InterPro: IPR000198; RhogAP.  
 DR InterPro: IPR000980; SH2.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00130; DAG\_PE\_bind; 1.  
 DR Pfam: PF00620; RhogAP; 1.  
 DR PRINTS: PR00008; DAGPEDOMAIN.  
 DR PRODOM: PD000093; SH2; 1.  
 DR SMART: SM00109; C1; 1.  
 DR SMART: SM00324; RhogAP; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
 DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM\_2; 1.  
 DR PROSITE: PS50001; SH2; 1.  
 DR KW: GTPase activation; Phorbol-ester binding; Zinc; SH2 domain;  
 KW: Alternative splicing.  
 FT DOMAIN 59 127 SH2.  
 FT DOMAIN 215 264 PHORBOL-ESTER AND DAG BINDING.  
 FT DOMAIN 291 435 RHO-GAP.  
 FT CONFLICT 1 6 MAASSN -> MRLL (IN REF. 1).  
 SQ SEQUENCE 468 AA; 53923 MW; 6325495805804C CRC64;

Query Match 9.1%; Score 293.5; DB 1; Length 468;  
 Best Local Similarity 25.7%; Pred. No. 1.4e-11;

Matches 99; Conservative 66; Mismatches 155; Indels 63; Gaps 12;

QY 199 KTRRSIGSAVDGNESTIVAKTVTPNDGPIEAVSTIETVP-----YTRSRKKTGT 252  
 118 KRESIHDLVTDGLITLYETKAA-----EYISKMTNPIYEHIGVATLLREKV--- 166  
 DB 253 QPNNSDSTLMSRLERP-----TETDSVGTPOSGM-----RLHDFPSKT 293  
 167 -----SRILSRSNPEPKTVTHEHTAVKISLVRAALTNHNDHFNTEKTHNFVHT 221  
 QY 294 VIKPESVPCGKRIRFGKLS--LKRDCRVVSHPECHDRCPPLICPIPLIGTVKIGGML 351  
 222 FRPFPNCEYCA-NFMGLIAGVRCSDGGLNVRKCKSKHVPNDGQPLKRIK-KVYCCDL 279  
 DB 352 ADVFSQTSPIPSIVVCHVEIDRGITETGLYRISGCDRTVVELKEKFLR-VKTVPLL 409  
 280 TTVLVKHNQRPVYVDICIREIARGLSKGLYRVSGETHIEVYKKAFFRDEKADISA 339  
 QY 410 SKVDDHAIGSLKDFLRNKEPLRFLNRARFMEAAETIDENSTIAAQANGELQOAN 469  
 340 NVYPDNIITIGALKLFRDLPVITYDYTSKTDAKISNADRELAHEVLMLEPPAH 399  
 DB 470 RDTLAFMLIHQRYAASPHTK-MDVANLAKVFGPTTVAHVVPNDPVTMSQDIKRPQV 528  
 400 YETLRVLMHLKRVTNMEKDNFMNAENLGIYFGPTLM--RPEDSTLTTLHDMRYOKLIY 457  
 QY 529 ERLISLPLEYWSQPMAYEONI 550  
 458 Q-----ILIEDV 466

RESULT 10  
 OPHL\_HUMAN STANDARD; PRT; 814 AA.  
 ID OPHL\_HUMAN  
 Q9UNM1; 075117; Q9UJ00; Q9BYS6; Q9BYS7;

DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Oligophrenin-1 like protein (GTPase regulator associated with focal  
 DE adhesion kinase).  
 GN OPHNL1 OR GRAF OR KIA00621.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta, Heart, and Liver;  
 RA Xia J.H., Tang X.X., Yu K.P., Pan Q., Dai H.P.;  
 RT "Molecular cloning of human oligophrenin-1 like (OPHNL1) gene,  
 RT complete cds.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), DISEASE, AND VARIANT LEUKEMIA SER-417.  
 RX MEDLINE=20381355; PubMed=10908648;  
 RA Borkhardt A., Bojesen S., Haas O.A., Fuchs U., Bartelheimer D.,  
 RA Loncarevic I.F., Bohle R.M., Harbott J., Repp R., Jaeger U.,  
 RA Viehmann S., Henn T., Korth P., Schair D., Lampert F.;  
 RT "The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and  
 RT both alleles are disrupted in three cases of myelodysplastic  
 RT syndrome/acute myeloid leukemia with a deletion 5q.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9168-9173(2000).  
 RN [3]  
 RP SEQUENCE OF 53-785 FROM N.A. (ISOFORMS 1 AND 2).  
 RA Bojesen S.E., Link C., Borkhardt A.;  
 RT "Genomic structure of the human GRAF gene.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 62-814 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: GTPase activating protein for Rho.  
 CC -1- SUBUNIT: Binds to the C-terminal of p125(FAK).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -1- DISEASE: A form of juvenile myelomonocytic leukemia is  
 CC characterized by a chromosomal translocation t(5;11)(q31;q23) that  
 CC involves OPHNL1 and MLL.  
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).





```

CC - FUNCTION: GTPASE-ACTIVATING PROTEIN FOR RAC AND CDC42. PROMOTES
CC THE EXCHANGE OF RAC OR CDC42-BOUND GDP BY GTP, THEREBY ACTIVATING
CC THEM.
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: HIGHLY ENRICHED IN THE BRAIN. MUCH WEAKER
CC EXPRESSION IN HEART, LUNG AND MUSCLE.
CC - SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
CC - SIMILARITY: CONTAINS 1 PH DOMAIN.
CC - SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC - SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC - SIMILARITY: STRONG, TO HUMAN BCR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL: 001147; AAC50063.1;
CC EMBL: L19704; AAC37519.1;
CC EMBL: L19705; AAC37518.1; ALT_INIT.
CC Genew; HGNC:BL; ABR.
CC MIM; 600365;
CC InterPro: IPR000008; C2.
CC InterPro: IPR001331; GDS_CDC24.
CC InterPro: IPR001849; PH.
CC InterPro: IPR00198; RHO-GAP.
CC InterPro: IPR00219; RHOGEF.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00620; RHO-GAP; 1.
CC Pfam; PF00621; RHOGEF; 1.
CC SMART; SM00233; C2; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00324; RHO-GAP; 1.
CC SMART; SM00325; RHOGEF; 1.
CC PROSITE; PS50004; C2_DOMAIN_2; 1.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS00741; DH_1; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC Guanine-nucleotide releasing factor; Alternative splicing.
CC
CC FT DOMAIN 91 284
CC FT DOMAIN 301 459
CC FT DOMAIN 463 595
CC FT DOMAIN 661 805
CC FT DOMAIN 417 420
CC FT VARSPLIC 1 82
CC
CC FT CONFLICT 67 67
CC FT CONFLICT 657 660
CC FT CONFLICT 761 761
CC FT SEQUENCE 859 AA; 97696 MW; 9FD50CD54FA9483 CRC64;
CC
CC Query Match 8.7%; Score 282; DB 1; Length 859;
CC Best Local Similarity 25.5%; Pred. No. 1.7e-10;
CC Matches 140; Conservative 67; Mismatches 201; Indels 140; Gaps 21;
CC
CC 48 DHELGKVKDLMKATERSSALDVKLNHRNQVDVETIKRORAEADCELENGIOIIREML 107
CC 374 DHELGKVKDLMKATERSSALDVKLNHRNQVDVETIKRORAEADCELENGIOIIREML 107
CC 108 MCDTSGSIOLSEOKSALAFI-----NRGQSSSNAGNRRLSTIDSGSILLDISPKNDE 163
CC 418 LLNLSPTIPRIHNRNGSYFLILSSDYRESEWRRAIKLQKDKDAVAVLSVEQLVIG 477
CC 164 SLMDSSLVKTFKLLKKRKRSTSRQFVDPGPGVKKTRISGSAVDGNEISIVAKTVTV 223
CC 478 S-----CFKLTAVNIPIVTSNKKDDDESPGLYGLFVHSAKFGQSANLYCTLEV 528

```

```

CC
CC 224 PNDGPIBAVSTIETVPWTRSRKRT-----GTLPD-WNS----- 257
CC 529 DSFG-----YFV-SKATRVFRDTAEKMBEETIELEGSOSILICYECY 574
CC
CC 258 DSTLSNR-----OLEPRT-ETDSVGTPOSGMGRHDFVSKTVIRESCV 301
CC 575 DKTKYKNKNNEIVDKIMKGQQLDPQVETFNHMT-----DIENMGIVVERSM 624
CC
CC 302 PCGKRIKFGKSLKCRDCRVVSHPECRCRCLPCLTPTLGPVKIGEGMLADFS-----Q 357
CC 625 KFTSR-----DMSLK-----RPSKQKOTGVFGKISVYTKR 655
CC
CC 358 TSPMISTIVHCVNEIEORGLTEGLYRISGCDRTVKEKEKEF-LRVTVPLSKVDDII 416
CC 656 ESKKVPYIVROCVEEYERKGIIEVGITRISGVAITDIALKAVFDANNKDILMSMDIN 715
CC
CC 417 AICSLKDFLRLNKEPLTLFRLNRAFMENAEITDEDNSIAAYQAVGELPQANRDTLAFI 476
CC 716 AIAGTLKLYFRELPEPLTLDRLYPAPFMGIALSDPAKENCMHLLRSLPDLNLTPLFL 775
CC
CC 477 MIHLORVQO-SPHTMDVANKAKVGPITV-----AAVAPNDPVTMSODIRQRY 527
CC 776 LEHLKRVAKKEPINKMSLNATVGPLLRPEVESKALHTSAAD--IWSHDVAQVQV 833
CC
CC 528 VERLSLP 535
CC 834 LLYYQHP 841
CC
CC
CC RESULT 12
CC YNS4_CAEEL STANDARD; PRT; 747 AA.
CC
CC AC P34588;
CC DT 01-FEB-1994 (Rel. 28, Created)
CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Hypothetical protein ZC21.4 in chromosome III.
CC GN ZC21.4
CC OS Caenorhabditis elegans.
CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
CC OC Rhabditidae; Pelodertinae; Caenorhabditis.
CC OX NCBI_TaxID=6239;
CC RN [1]
CC RP SEQUENCE FROM N. A.
CC RC STRAIN-Bristol N2;
CC RX MEDLINE=94150718; PubMed=7906398;
CC RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
CC RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
CC RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
CC RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
CC RA Johnston L., Jones M., Kerhaw J., Kirsten J., Lalister N.,
CC RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
CC RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
CC RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
CC RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
CC RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
CC RA Wohldman P.;
CC RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
CC RT elegans."
CC RL Nature 368:32-38(1994).
CC CC - SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL: L16685; AAA28171.1; ALT_INIT.
CC HSP; 007960; IRCP.

```

DR Wormpep: ZC21.4; CE00534.  
 DR InterPro: IPR000198; RhogAP.  
 DR Pfam: PF00620; RhogAP; 1.  
 DR SMART: SM00324; RhogAP; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 747 AA; 80829 MW; 48E9D61EDC4C56B0 CRC64;  
 Query Match 8.6%; Score 280.5; DB 1; Length 747;  
 Best Local Similarity 25.1%; Pred. No. 1.7e-10;  
 Matches 105; Conservative 59; Mismatches 146; Indels 109; Gaps 14;  
 QY 223 VPDGSPPIEAVSTIEFVPTWTSRRKTYGLQPMNSDSTLSROLEPPTETDVGTPQSNQ 292  
 DB 35 IENGEPPTATATPKSGRKWKSKAKOGSGSGSSSSSSGSOO-----OGAAGAPQPV 88  
 QY 283 GMLHDFVSKTVIKPSSVCPCGRKIRFGKSLKRCRCRVSHPECDRCRLPCLPIIGT 342  
 DB 89 GVAIAD-----CPTG-----SCEDH----- 103  
 QY 343 PVKIGGMLADPVSTSPMIPSIYVHCVEIEQRLTGLYRISGCDRTVKELKEK--- 399  
 DB 104 -----VPMIVQACVCYETETGMDTVGIRIPGTAIVNALKESLN 144  
 QY 400 --FLAKTYVPLSKVD-----DIAICSLKDFLRNKEPLITRLNFAEMAEITD 450  
 DB 145 RGEDSDV-----LSKVESLDPRMDDVNVSSILKMFRLKLPPELITLKYFPFDANRIST 200  
 QY 451 EDNSIAAMTQAVGELQVANOADLAFMIHQLQVAVO--SPHKMDVANLAKVFGPTIYAHV 509  
 DB 201 HHNRKLKRLNKLRLKLRPHRPHDITLRFILVHLSSETKSDVKMKCRNLALMEGSIYR--- 257  
 QY 510 PNDPV--TMSQDIKRPKVERLLSLPLEYWSQFMVDEIDPLHYENSNAFSPQPT 568  
 DB 258 PSDDNATGWTHTMSDCKIETETLIHYNLMMFDESSPTE--DAVEQHPADQGN---PLEP 312  
 QY 569 -----DIKYSILGPTVTPPEHQLKTPSSSSLS-----QRRSRILTKTKPPFGSK 614  
 DB 313 GGTCVGPPTGVSAASFNDMHLIRKANEDQAAAMNEGKQKIKMLRNSRDKSKSK 371  
 RESULT 13  
 BCR\_HUMAN STANDARD; PRT; 1271 AA.  
 AC P1274; Q12842;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Breakpoint cluster region protein (BC 2.7.1.-).  
 GN BCR OR BCR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RX MEDLINE=95394474; PubMed=7665185;  
 RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burtian D.,  
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jhan L., Ma Y.,  
 RA McLaurin H.-J., Pan H.-O., Sarhan O.H., Toth S., Wang Z., Zhang G.,  
 RA Heisterkamp N., Groffen J., Roe B.A.;  
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
 RT regions involved in the Philadelphia chromosomal translocation.";  
 RL Genomics 27:67-82(1995).  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87218455; PubMed=3107980;  
 RA Hariharan I.K., Adams J.M.;  
 RT "cDNA sequence for human bcr, the gene that translocates to the abl  
 RT oncogene in chronic myeloid leukaemia.";  
 RL EMBO J. 6:115-119(1987).  
 RN (3)  
 RP SEQUENCE OF 683-1271 FROM N.A.  
 RX MEDLINE=85240564; PubMed=2989703;

RA Heisterkamp N., Stam K., Groffen J., de Klein A., Grosveld G.;  
 RT "Structural organization of the bcr gene and its role in the Ph"  
 RT translocation";  
 RL Nature 315:758-761(1985).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88217290; PubMed=3285291;  
 RA Lifshitz B., Feinstein E., Marcelle C., Shlivelman E., Amson R.,  
 RA Gale R.P., Canaan E.;  
 RT "bcr genes and transcripts.";  
 RL Oncogene 2:113-117(1988).  
 RN (5)  
 RP SEQUENCE OF 1-693 FROM N.A.  
 RX MEDLINE=87092329; PubMed=3540951;  
 RA Mes-Masson A.M., McLaughlin J., Daley G.O., Paskind M., Witte O.N.;  
 RT "Overlapping cDNA clones define the complete coding region for the  
 RT p210-abl gene product associated with chronic myelogenous leukemia  
 RT cells containing the Philadelphia chromosome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9768-9772(1986).  
 RN (6)  
 RP REVISIONS.  
 RA Mes-Masson A.M., McLaughlin J., Daley G.O., Paskind M., Witte O.N.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2507-2507(1987).  
 RN (7)  
 RP SEQUENCE OF 1-46 AND 275-426 FROM N.A.  
 RX MEDLINE=91088292; PubMed=2263470;  
 RA Zhu O.S., Heisterkamp N., Groffen J.;  
 RT "Unique organization of the human BCR gene promoter.";  
 RL Nucleic Acids Res. 18:7119-7125(1990).  
 RN (8)  
 RP SEQUENCE OF 56-426 FROM N.A.  
 RX MEDLINE=88065859; PubMed=2825022;  
 RA Feinstein E., Marcelle C., Rosner A., Canaan E., Gale R.P.,  
 RA Dreazen O., Smith S.D., Croce C.M.;  
 RT "A new fused transcript in Philadelphia chromosome positive acute  
 RT lymphocytic leukemia";  
 RL Nature 330:386-388(1987).  
 RN (9)  
 RP SEQUENCE OF 1-4 FROM N.A.  
 RX MEDLINE=91172169; PubMed=1900918;  
 RA Shah N.P., Witte O.N., Denny C.T.;  
 RT "Characterization of the BCR promoter in Philadelphia chromosome-  
 RT positive and negative cell lines";  
 RL Mol. Cell. Biol. 11:1854-1860(1991).  
 RN (10)  
 RP FUNCTION.  
 RX MEDLINE=91238969; PubMed=1903516;  
 RA Diekmann D., Brill S., Garrett M.D., Toty N., Hsuan J.,  
 RA Montfries C., Hall C., Lim L., Hall A.;  
 RT "Bcr encodes a GTPase-activating protein for p21rac.";  
 RL Nature 351:400-402(1991).  
 RN (11)  
 RP BINDING TO ABL SH2-DOMAIN.  
 RX MEDLINE=91300547; PubMed=1712671;  
 RA Pendergast A.M., Muller A.J., Havlik M.H., Maru Y., Witte O.N.;  
 RT "BCR sequences essential for transformation by the BCR-ABL oncogene  
 RT bind to the ABL SH2 regulatory domain in a non-phosphotyrosine-  
 RT dependent manner.";  
 RL Cell 66:161-171(1991).  
 RN (12)  
 RP KINASE ACTIVITY.  
 RX MEDLINE=92034969; PubMed=1657398;  
 RA Maru Y., Witte O.N.;  
 RT "The BCR gene encodes a novel serine/threonine kinase activity within  
 RT a single exon.";  
 RL Cell 67:459-468(1991).  
 CC -/- FUNCTION: GTPASE-ACTIVATING PROTEIN FOR RAC1 AND CDC42. PROMOTES  
 CC THE EXCHANGE OF RAC OR CDC42-BOUND GDP BY GTP, THEREBY ACTIVATING  
 CC THEM. DISPLAYS SERINE/THREONINE KINASE ACTIVITY.  
 CC -/- DOMAIN: THE REGION INVOLVED IN BINDING TO ABL SH2-DOMAIN IS RICH  
 CC IN SERINE RESIDUES AND NEEDS TO BE SER/THR PHOSPHORYLATED PRIOR TO  
 CC SH2 BINDING. THIS REGION IS ESSENTIAL FOR THE ACTIVATION OF THE  
 CC ABL TYROSINE KINASE AND TRANSFORMING POTENTIAL OF THE CHIMERIC

CC BCR-ABL ONCOGENE.  
 CC - P.TM: AUTOHOMOPHORYLATED.  
 CC - DISEASE: Participates in a t(9;22)(q34;q11) chromosomal  
 CC translocation that produces a BCR-ABL oncogene responsible for  
 CC chronic myeloid leukemia (CML), acute myeloid leukemia (AML) and  
 CC acute lymphoblastic leukemia (ALL).  
 CC - SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.  
 CC - SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC - SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 CC - SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.  
 CC - SIMILARITY: STRONG, TO HUMAN ABR AND DROSOPHILA ROTUND.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: U07000; AAB60388.1; -  
 CC EMBL: M24603; AAB65939.1; -  
 CC EMBL: Y00661; CAA68676.1; -  
 CC EMBL: X02596; CAA26441.1; -  
 CC EMBL: M15025; AAA35594.1; -  
 CC EMBL: X52828; CAA37010.1; -  
 CC EMBL: X52829; CAA37011.1; -  
 CC EMBL: M64437; -; NOT\_ANNOTATED\_CDS.  
 CC PIR: A26172; TVHUBR.  
 CC PIR: A26664; TVHUA2.  
 CC PIR: A28765; TVHUA3.  
 CC Genew: HGNC:1014; BCR.  
 CC MIM: 151410; -;  
 CC  
 CC InterPro: IPR000008; C2.  
 CC InterPro: IPR001331; GDS\_CDC24.  
 CC InterPro: IPR001849; PH.  
 CC InterPro: IPR000198; RhogAP.  
 CC InterPro: IPR000219; RhogEF.  
 CC Pfam: PF00168; C2; 1.  
 CC Pfam: PF00169; PH; 1.  
 CC Pfam: PF00620; RhogAP; 1.  
 CC Pfam: PF00621; RhogEF; 1.  
 CC SMART: SM00233; C2; 1.  
 CC SMART: SM00233; PH; 1.  
 CC SMART: SM00324; RhogAP; 1.  
 CC SMART: SM00325; RhogEF; 1.  
 CC PROSITE: PS50004; C2\_DOMAIN\_2; 1.  
 CC PROSITE: PS50010; DH\_2; 1.  
 CC PROSITE: PS50041; DH\_1; 1.  
 CC PROSITE: PS50003; PH\_DOMAIN; 1.  
 CC Guanine-nucleotide releasing factor; Proto-oncogene; Transferase;  
 CC Chromosomal translocation; Serine/threonine-protein kinase;  
 CC Phosphorylation.  
 CC  
 CC FT DOMAIN 1 426 KINASE.  
 CC FT DOMAIN 498 691 DH.  
 CC FT DOMAIN 708 866 PH.  
 CC FT DOMAIN 870 1002 C2 DOMAIN.  
 CC FT DOMAIN 1068 1212 RHO-GAP.  
 CC FT BINDING 197 365 TO ABL SH2-DOMAIN.  
 CC FT DOMAIN 824 827 POLY-LEU.  
 CC FT SITE 426 427 BREAKPOINT FOR TRANSLLOCATION TO FORM  
 CC BCR-ABL ONCOGENE.  
 CC  
 CC FT CONFLICT 287 287 M -> I (IN REF. 4).  
 CC FT CONFLICT 418 418 G -> D (IN REF. 4).  
 CC FT CONFLICT 483 483 E -> K (IN REF. 4).  
 CC FT CONFLICT 560 560 E -> S (IN REF. 4).  
 CC FT CONFLICT 733 733 E -> D (IN REF. 1 AND 4).  
 CC FT CONFLICT 796 796 S -> N (IN REF. 1).  
 CC FT CONFLICT 961 1004 MISSING (IN REF. 4).  
 CC SEQUENCE 1271 AA; 142805 MW; A59F76B6F1AF3236 CRC64;

Query Match 8.5%; Score 275.5; DB 1; Length 1271;  
 Best Local Similarity 24.0%; Pred. No. 7.5e-10;

Matches 137; Conservative 86; Mismatches 196; Indels 153; Gaps 28;  
 QY 43 KW-----QRTHELGLKYDILMKATKTESALDYKIKHARNOYDVIRKRRQREAC 93  
 DB 754 KWTPLTDSFQWVP-ELFVAVNPITLVPPEELDAKIKI-----SQKSDIOEKRANKGS 808  
 QY 94 EKLEROIQLIRELMKODTSGSIQLSSEKSAFLNRGQPSSSNAGKRLSTIDEGSTL 153  
 DB 809 KATER-----LKKKL-----SEQSLILLSPSNAFVHSHNGSYTF-----LIS 849  
 QY 154 SDISEPKTDESIDWSSLVKTKRKRKRSTSRQFVDPGPGVKKRTSGSAVDQ--G 211  
 DB 850 SDY-----ERAEMRENI-----REQQKCFRSF-----SLTSVELQMLT 883  
 QY 212 NESIYAKTVYVP-----NDGSP-----LEAVSTIEYVYVW-- 242  
 DB 884 NSCVKLQYVHSHIPITLNKEDDSPGLYGLNVIHSAIGFKOSSNLYCTLEVDSEYFVN 943  
 QY 243 -TRSRKKTGTLP--WNSDSTL--NSROLE-----PRTETDVGTPQSGMG 284  
 DB 944 KAKTRYRTAEPNWNNEFEIELESGQTLRIICEKYNKTKIPKEDGSTRLMKQGV 1003  
 QY 285 RL-----HDFVSKTVIKPESCVPGCKRIKFG--KLSKCDRCRVVSHPECDRCPLPC 335  
 DB 1004 QLDPALQDRDW-QRTVIAMNG-IEVKLSVKFNSREFSLK-----RMPSRK 1047  
 QY 336 IPLLIGTPVKGIGMADVVSOTSPMIPSIYVHCVEIQRGILTEGLTRISGCDRTVE 395  
 DB 1048 QTVGEG--VKI-----AVYTKRERSKVPYIVRCVBEIERGMEEGIVRVSATVDIOA 1100  
 QY 396 LKEKF-LRKYTVPLSKVDIHAICSLDPLRNKEPLTLPLRNAPFEAAETPEDNS 454  
 DB 1101 LKAAPVNNKNDVSVAMSEMDVNAIAGTLLTYRELPELFTDEFYVNFEGIALSPVAK 1160  
 QY 455 IAATQAVGELPQANDTLAFLMILQVAVSPH--TKMDVANIAYKGGPTV-----AH 507  
 DB 1161 ESCMNLILSLPEANILFLFLDLHLKRAVEKAVKMSLHNLATVFGPTLLRPSRESK 1220  
 QY 508 AVPNP-DPYTM-----SQDLKRPKYVERLSL 534  
 DB 1221 LPANPSQPTMTDMSLEVMQOVLLYPLQL 1252

RESULT 14  
 OPRL\_HUMAN  
 ID OPRL\_HUMAN STANDARD; PRT; 802 AA.  
 AC 060890; Q8WX47;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Oligophrenin 1.  
 GN OPHNL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND DISEASE.  
 RC TISSUE-Fetal Brain;  
 RX MEDLINE=98241228; PubMed=9582072;  
 RA Billaut P., Bievenue T., Ronce N., des Portes V., Vinet M.C.,  
 RA Zemi R., Roest Crolius H., Carrie A., Fauchereau F., Cherry M.,  
 RA Billaut S., Hamel B., Frys J.-P., Beldjord C., Kahn A., Moraine C.,  
 RA Chelly J.;  
 RA "Oligophrenin-1 encodes a rhogap protein involved in x-linked mental  
 RA retardation.";  
 RL Nature 392:923-926(1998).  
 RP  
 RP SEQUENCE FROM N.A., AND VARIANTS THR-45 AND MET-301.  
 RX PubMed=10818214;  
 RA Billuart P., Chelly J., Carrie A., Vinet M.C., Couvert P.,  
 RA McDonnell N., Zemi R., Kahn A., Moraine C., Beldjord C., Bievenue T.;  
 RA "Determination of the gene structure of human oligophrenin-1 and







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:56:22 ; Search time 55 Seconds

(without alignments)  
1321.373 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMLNVNLEFQVLRVE.....SKSATNIGRGNFASPMUK 632

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 11492915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3243	100.0	632	10	US-09-881-736-2
2	3230	99.6	632	10	US-09-883-790-413
3	2919	90.0	570	10	US-09-883-790-427
4	2734	84.3	628	10	US-09-881-736-4
5	541.5	16.7	681	10	US-09-881-736-6
6	294.5	9.1	193	10	US-09-802-127-8
7	286.5	8.8	1286	9	US-10-153-668-212
8	286.5	8.3	1354	9	US-10-153-668-470
9	270.5	8.3	2548	10	US-09-851-682A-1
10	263.5	8.1	555	9	US-09-764-868-878
11	251	7.7	170	9	US-10-080-960-32
12	251	7.7	170	10	US-09-802-127-7
13	247.5	7.6	291	9	US-09-764-868-898
14	223.5	6.9	152	9	US-10-132-585-5
15	218	6.7	103	9	US-10-080-960-34
16	213.5	6.6	429	9	US-09-764-868-1029
17	210.5	6.5	803	10	US-09-770-689A-2
18	206.5	6.4	257	9	US-09-764-868-1024
19	205.5	6.3	881	10	US-09-816-860A-2

20	205	6.3	726	10	US-09-770-689A-4	Sequence 4, Appli
21	204.5	6.3	780	10	US-09-770-689A-5	Sequence 5, Appli
22	200.5	6.2	1702	9	US-09-854-133-434	Sequence 434, App
23	200.5	6.2	1702	9	US-10-144-649A-434	Sequence 434, App
24	200.5	6.2	1702	10	US-09-738-973-434	Sequence 434, App
25	198.5	6.1	254	9	US-09-764-868-1014	Sequence 1014, Ap
26	197.5	6.1	803	9	US-10-059-585-10	Sequence 10, Appl
27	194	6.0	731	9	US-10-201-481-3	Sequence 3, Appl
28	193.5	6.0	643	9	US-10-097-340-16	Sequence 16, Appl
29	192	5.9	390	9	US-10-097-340-21	Sequence 21, Appl
30	192	5.9	464	9	US-10-097-340-14	Sequence 14, Appl
31	192	5.9	718	9	US-10-097-340-19	Sequence 19, Appl
32	192	5.9	751	9	US-10-097-340-18	Sequence 18, Appl
33	189	5.8	663	9	US-10-080-960-14	Sequence 14, Appl
34	187	5.8	2167	10	US-09-801-368-56	Sequence 56, Appl
35	186.5	5.8	341	9	US-09-764-868-1041	Sequence 1041, Ap
36	179.5	5.5	327	9	US-09-764-868-631	Sequence 631, App
37	179.5	5.5	327	9	US-09-955-999-74	Sequence 74, Appl
38	174.5	5.4	547	10	US-09-802-127-2	Sequence 2, Appli
39	164	5.1	332	9	US-09-764-868-1040	Sequence 1040, Ap
40	160.5	4.9	332	9	US-09-764-868-1225	Sequence 1225, Ap
41	159.5	4.9	188	10	US-09-925-300-1017	Sequence 1017, Ap
42	156.5	4.8	288	9	US-09-764-868-1048	Sequence 1048, Ap
43	154.5	4.8	162	9	US-09-764-868-656	Sequence 656, App
44	148.5	4.6	2099	9	US-10-128-714-3290	Sequence 3290, Ap
45	148.5	4.6	2405	9	US-10-128-714-8250	Sequence 8250, Ap

## ALIGNMENTS

RESULT 1  
US-09-881-736-2  
; Sequence 2, Application US/09881736  
; Patent No. US20020076785A1  
; GENERAL INFORMATION:  
; APPLICANT: Glotzer, Michael  
; APPLICANT: Jantsch-Plunger, Verena  
; APPLICANT: Romano, Alper  
; APPLICANT: Mishima, Masanori  
; APPLICANT: Kaitna, Susanne  
; TITLE OF INVENTION: Cyt-4 polypeptides, DNA molecules encoding them and their use  
; FILE REFERENCE: 0652.226001/EKS/AES  
; CURRENT APPLICATION NUMBER: US/09/881,736  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: EP 00 112 880.0  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: EP 01 110 554.1  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/241,231  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: To be determined  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 632  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-881-736-2

Query Match 100.0%; Score 3243; DB 10; Length 632;  
Best Local Similarity 100.0%; Pred. No. 2.7e-239;  
Matches 632; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTMLNVNLEFQVLRVEISSEGEVFOILADFEFRKKMORTHELKRYKDLMLK 60  
DB 1 MDTMLNVNLEFQVLRVEISSEGEVFOILADFEFRKKMORTHELKRYKDLMLK 60  
QY 61 AETRSALDVKLKHANQVDEIKRRORAEADCEKLEKRIQILIRELMKDTSGSIQLSSE 120  
DB 61 AETRSALDVKLKHANQVDEIKRRORAEADCEKLEKRIQILIRELMKDTSGSIQLSSE 120

QY	121	OKSALAFINRQGPSSSMNGKRLSTIDEGSILSDISFPKTBESLMDWSLYVTFKLKR	180
Db	121	OKSALAFINRQGPSSSMNGKRLSTIDEGSILSDISFPKTBESLMDWSLYVTFKLKR	180
QY	181	EKRKSTSQFDGPPGPAVKTRTSGSAVDQGNESIYAKTTVTVPNDGPFLEAVSTIETVP	240
Db	181	EKRKSTSQFDGPPGPAVKTRTSGSAVDQGNESIYAKTTVTVPNDGPFLEAVSTIETVP	240
QY	241	YWTBRRRKRTGTLQPMNSDSTLSNROLEPTEITDVSQTPQSNQGMRLHDEVSQVIRKEPSC	300
Db	241	YWTBRRRKRTGTLQPMNSDSTLSNROLEPTEITDVSQTPQSNQGMRLHDEVSQVIRKEPSC	300
QY	301	VPCGKRIFGKLSLKCRCDCRVVSHPECRCBPLCIPITLIGTPVKIGEGMLADFVSQTSB	360
Db	301	VPCGKRIFGKLSLKCRCDCRVVSHPECRCBPLCIPITLIGTPVKIGEGMLADFVSQTSB	360
QY	361	MIPSTIVYHCVNIIEORGLTEGTLNRSQGDRTVYKELKEFLRKTYPLLSKYNDTHAITS	420
Db	361	MIPSTIVYHCVNIIEORGLTEGTLNRSQGDRTVYKELKEFLRKTYPLLSKYNDTHAITS	420
QY	421	LLKDELRLNLKEPFLFRLNRAFMEEAEITDEDNSIAAMQAVGELPQANRDTLAFIMIH	480
Db	421	LLKDELRLNLKEPFLFRLNRAFMEEAEITDEDNSIAAMQAVGELPQANRDTLAFIMIH	480
QY	481	QRVASOPTKMDVANIYAKYFEGPTIYAHANVPNDPVTMSODIKRQKRYVERLISLPLEYMS	540
Db	481	QRVASOPTKMDVANIYAKYFEGPTIYAHANVPNDPVTMSODIKRQKRYVERLISLPLEYMS	540
QY	541	QFMAYEQRNIDPLAHYIENSNAFSTPQTPDIKLSILGPAVTPPEHQLLKTPESSSLSORVMS	600
Db	541	QFMAYEQRNIDPLAHYIENSNAFSTPQTPDIKLSILGPAVTPPEHQLLKTPESSSLSORVMS	600
QY	601	TLTKRTPPRGSKSKSATINLGRQGNFPASPMK	632
Db	601	TLTKRTPPRGSKSKSATINLGRQGNFPASPMK	632

```

      RESULT 2
      US-09-833-790-413
      ; Sequence 413, Application US/09833790
      ; Patent No. US2002006288A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Lodes, Michael J.
      ; APPLICANT: Wang, Tonglong
      ; APPLICANT: Secretist, Heather
      ; APPLICANT: Mohamach, Raodoh
      ; APPLICANT: Inditilas, Carol Y.
      ; APPLICANT: Fan, Liqun
      ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
      ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
      ; FILE REFERENCE: 210121.512
      ; CURRENT APPLICATION NUMBER: US/09/833,790
      ; CURRENT FILING DATE: 2001-04-11
      ; NUMBER OF SEQ ID NOS: 440
      ; SOFTWARE: FastSeq for Windows Version 4.0
      ; SEQ ID NO 413
      ; LENGTH: 632
      ; TYPE: PRT
      ; ORGANISM: Homo sapiens
      US-09-833-790-413

      Query Match          99.6%;      Score 3230;      DB 10;      Length 632;
      Best Local Similarity 99.7%;      Pred. No. 2,6e+236f;
      Matches 630;      Conservative 0;      Mismatches 2;      Indels 0;      Gaps 0
    
```

QY	121	OKSALAFINRGQPPSSNMGNKRLLSTIDESGSLISDIPSPKTPESLDIMOSLKYTRFLKKR	180
Db	121	OKSALAFINRGQPPSSNMGNKRLLSTIDESGSLISHSI SEDKTDIESLDIMOSLKYTRFLKKR	180
QY	181	EKRSTSHQFVDPGPPVKKTRTSGSAVDQGNESIVAKTIVVPNDGPIEAVSTIETVP	240
Db	181	EKRSTSHQFVDPGPPVKKTRTSGISAVDQGNESIVAKTIVVPNDGPIEAVSTIETVP	240
QY	241	YWT8RRRTGTLOPNNSDSTLNSRLEPTEITDVGTPQNSNGMRLHDPVSTAVIKPESC	300
Db	241	YWT8RRRTGTLOPNNSDSTLNSRLEPTEITDVGTPQNSNGMRLHDPVSTAVIKPESC	300
QY	301	VPCGKRIFKGSLILKCRDQVWVSHDECBDRCPLPIPTLIGPVKIGEGMLADFEVSQTS	360
Db	301	VPCGKRIFKGSLILKCRDQVWVSHDECBDRCPLPIPTLIGPVKIGEGMLADFEVSQTS	360
QY	361	MIPSTIVHCYVNIIEORGLETEGLYRISGCDRIVKELKEFLKAVKVPYLSTKYDILHACS	420
Db	361	MIPSTIVHCYVNIIEORGLETEGLYRISGCDRIVKELKEFLKAVKVPYLSTKYDILHACS	420
QY	421	LKLDPLRLMKBELILFFRLNRAFMEEAETDEDNESIAMYAQAVGELPQANRDYLAFLMIHL	480
Db	421	LKLDPLRLMKBELILFFRLNRAFMEEAETDEDNESIAMYAQAVGELPQANRDYLAFLMIHL	480
QY	481	ORVASOPTHKDVANLAKVFGPTIYAAHAVPNNDPYTMSODIKRQKRYVERLLSTPLEYVS	540
Db	481	ORVASOPTHKDVANLAKVFGPTIYAAHAVPNNDPYTMSODIKRQKRYVERLLSTPLEYVS	540
QY	541	QPMAYEONIDPLAHYIENSNAFSTQPTDIKYSLLGAPVYTPPEHOLKLTPTSSSSLSQRYVS	600
Db	541	QPMAYEONIDPLAHYIENSNAFSTQPTDIKYSLLGAPVYTPPEHOLKLTPTSSSSLSQRYVS	600
QY	601	TLTKNTPPRGSKSATSATNLGROGNFPASPMUK 632	
Db	601	TLTKNTPPRGSKSATSATNLGROGNFPASPMUK 632	

```

RESULT 3
US-09-833-790-427
; Sequence 427, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Momahach, Raodoh
; APPLICANT: Indriasis, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-427

Query Match          90.0%;      Score 2919;  DB 10;      Length 570;
Best Local Similarity 99.8%;      Pred. No. 1.3e+214;
Matches 569;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```



QY	183	RRSRSTRQVDDGPGGVKTRSTRISGSAVDGONESTIAKTTVTYPNDGEPITAVSTIEVPMW	242
Db	121	RRSTRSQVDDGPGGVKTRSTRISGSAVDGONESTIAKTTVTYPNDGEPITAVSTIEVPMW	180
QY	243	TRSRKKTGTLPWNDSSTLNSHOLEPRTETDVSQVTSQNGKRLHDFVSKTVYKPPSCVW	302
Db	181	TRSRKKTGTLPWNDSSTLNSHOLEPRTETDVSQVTSQNGKRLHDFVSKTVYKPPSCVW	240
QY	303	CGKRLKFKKLSLKCHDCRVSHPEBCRDRCPLPILSTIPVKIGSGMLADYVSQISPMI	362
Db	241	CGKRLKFKKLSLKCHDCRVSHPEBCRDRCPLPILSTIPVKIGSGMLADYVSQISPMI	300
QY	363	PSIVVHCNEIEORLTETGLYRISGCDRTYKELKEFLRVKTVPLLSKYVDIHAICSLI	422
Db	301	PSIVVHCNEIEORLTETGLYRISGCDRTYKELKEFLRVKTVPLLSKYVDIHAICSLI	360
QY	423	KDPLRLNKEPLTLTFLRLNRAFMFAEITDEDSIAAMYQAVGELPQANRODTLAFMLTHLOR	482
Db	361	KDPLRLNKEPLTLTFLRLNRAFMFAEITDEDSIAAMYQAVGELPQANRODTLAFMLTHLOR	420
QY	483	VAQSPHTMDVANAIAKVFGRPTIVAAYVPRDVTMSODIKRQPKVYERLISLPLEYWSQF	542
Db	421	VAQSPHTMDVANAIAKVFGRPTIVAAYVPRDVTMSODIKRQPKVYERLISLPLEYWSQF	480
QY	543	MMVEQENIDPLPHVIENSNAFSTPQTPDIKVSLLGVPYTPBEHOLKTPSSSSLSQVRSTL	602
Db	481	MMVEQENIDPLPHVIENSNAFSTPQTPDIKVSLLGVPYTPBEHOLKTPSSSSLSQVRSTL	540
QY	603	TKNTPRSGSKSANTLGRQGFASFPMUK	632
Db	541	TKNTPRSGSKSANTLGRQGFASFPMUK	570

```

RESULT 4
US-09-881-736-4
; Sequence 4, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kalina, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use in
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/ABS
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-881-736-4

```

Query Match      84.3%    Score 2734; DB 10; Length 628;  
Best Local Similarity    84.4%    Pred. No. 2e-200;  
Matches    534; Conservative    47; Mismatches    46; Indels    6; Gaps    3.

OY    1 MDTMMLVNRFEDLYRRVELSGCNF-VQITOLAKDEEDRKKMQRPDHDLGKYKIDLL 59  
         | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB    1 MDTVMDLMTLEFDLYRRMELIINGNSIEFTIQVADKEEDRKKRYQRTINDELKRDILL 60

[illegible]

```

RESULT 5
US-09-881-736-6
; Sequence 6, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Piunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kalina, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/ABS
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-881-736-6

```

Query Match 16.7%; Score 541.5; DB 10; Length 681;



```

Db 354 DMLAMIKTIOESSNLNE-EDGVTNRDLIS---RRIKEYNMLMSKAEQ-----LPKT-- 402
Qy 222 TVPNDGPLEAVSTIETVYPTWRSRRKTGTLOPMNSDSTLNSRQLEPRTETDSVGTPOSN 281
Db 403 -----PROSLIQTLL-LGAKSEPKT---OSPHSPKEESERKL---LSKDDTSPPKDK 448
Qy 282 GGMRL-LHDFVSKTV-IKPESCVPGKRIKFKLSLKCDKCVVSHPEKDRCPJPCIPPT 338
Db 449 GWRKGIPSIIMKTEKKPTATGTGVRLL-----DDCP----- 481
Qy 339 LIGTFVKIGEGMLADVFVSTSPMIPSIYVHCYNEIEORGLTETGLYRISGCRITYKELKE 398
Db 482 -----PAHTNRYIPLVIDICKLVEERLEETGTGYRVPGNNAATSSMQE 525
Qy 399 KFLRVKTVPLSKVD-----DHAICSLKDFLRLNKEPLTFRLNRAFEAAETIDE 451
Db 526 ELNKG-----MADIDIODDKWDLNVISLSKSFRRKLPEPLFTNDKYADFEANRKEBP 580
Qy 452 DNSIAAMTQAVGELPOANDTLAFIMHLORYAQ-SPHTKMDVANLAKVFGPTIAHAMP 510
Db 581 LDRLTFLRLHDLPEHHEITKFLSAHLKTVAEENSEKKNMPEPRNLAIYFGPTLVRTSBD 640
Qy 511 NPDPTMSODIKROP---KVERLLSLPLEYWSQFMAYEQENIDPLHVIENSNASTPQT 567
Db 641 N-----MTHVTHMPDQRYIVETLIQ-----HDMFTEEGAEPLTVQEESTVDSQPV 690
Qy 568 PDIKVSLGPTVTPPEHOLKT---PSSSLSQVRSTLTKTPTPRGS 611
Db 691 PNI-----DHLTLNIGRTGVSPDVSATSSTKSGSMGS 727

```

## RESULT 8

```

US-10-153-668-470
; Sequence 470, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 470
; LENGTH: 1354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-470

```

```

Query Match 8.8%; Score 286.5; DB 9; Length 1354;
Best Local Similarity 22.1%; Pred. No. 5.2e-13;
Matches 143; Conservative 98; Mismatches 216; Indels 191; Gaps 26;

```

```

Qy 48 DHE-LGKYK-DILMAETERSAL-----DVKL-KHARNOVVEIKRRQRAEADC 93
Db 255 DHEVGPSPSLDAQPQSTERSKSYDEGLDDYREDAKLSFKHVSLSKIKIADSQSSSDS 314

```

```

Qy 94 -----EKTERIOIOLIRELMQDTSIGLSSEOKS 123
Db 315 GSRKSSSEVEFSDAKKEGHLHFRPLVTDKGGKVGSSIRWKOMYVYLKSHSLYLKDKRE 374
Qy 124 ALAPLNRGQSSSSNAGNKRLLSTIDSGSLSDISDPKT----- 161
Db 375 QTPSEEBEQPISVNA-----CLIDISYSETRKKNVFLRTTSDCECLPQAEOR 421
Qy 162 DESLMDSSLVYKTFELKKRKRRTSRQFVDPGPVKTKRSIGSADQNGESIYAKTTV 221
Db 422 DMLAMIKTIOESSNLNE-EDGVTNRDLIS---RRIKEYNMLMSKAEQ-----LPKT-- 470
Qy 222 TVPNDGPLEAVSTIETVYPTWRSRRKTGTLOPMNSDSTLNSRQLEPRTETDSVGTPOSN 281
Db 471 -----PROSLIQTLL-LGAKSEPKT---OSPHSPKEESERKL---LSKDDTSPPKDK 516
Qy 282 GGMRL-LHDFVSKTV-IKPESCVPGKRIKFKLSLKCDKCVVSHPEKDRCPJPCIPPT 338
Db 517 GWRKGIPSIIMKTEKKPTATGTGVRLL-----DDCP----- 549
Qy 339 LIGTFVKIGEGMLADVFVSTSPMIPSIYVHCYNEIEORGLTETGLYRISGCRITYKELKE 398
Db 550 -----PAHTNRYIPLVIDICKLVEERLEETGTGYRVPGNNAATSSMQE 593
Qy 399 KFLRVKTVPLSKVD-----DHAICSLKDFLRLNKEPLTFRLNRAFEAAETIDE 451
Db 594 ELNKG-----MADIDIODDKWDLNVISLSKSFRRKLPEPLFTNDKYADFEANRKEBP 648
Qy 452 DNSIAAMTQAVGELPOANDTLAFIMHLORYAQ-SPHTKMDVANLAKVFGPTIAHAMP 510
Db 649 LDRLTFLRLHDLPEHHEITKFLSAHLKTVAEENSEKKNMPEPRNLAIYFGPTLVRTSBD 708
Qy 511 NPDPTMSODIKROP---KVERLLSLPLEYWSQFMAYEQENIDPLHVIENSNASTPQT 567
Db 709 N-----MTHVTHMPDQRYIVETLIQ-----HDMFTEEGAEPLTVQEESTVDSQPV 758
Qy 568 PDIKVSLGPTVTPPEHOLKT---PSSSLSQVRSTLTKTPTPRGS 611
Db 759 PNI-----DHLTLNIGRTGVSPDVSATSSTKSGSMGS 795

```

## RESULT 9

```

US-09-851-682A-1
; Sequence 1, Application US/09851682A
; Patent No. US20020091248A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/851,682A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: US/09/172,422
; PRIOR FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-851-682A-1

```

```

Query Match 8.3%; Score 270.5; DB 10; Length 2548;
Best Local Similarity 24.4%; Pred. No. 2.2e-11;
Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

```



```

: CURRENT APPLICATION NUMBER: US/09/802,127
: CURRENT FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185,611
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 170
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Rho-Gap consensus sequence
: US-09-802-127-7

```

Query Match	7.7%;	Score 251;	DB 10;	Length 170;
Best Local Similarity	39.8%;	Pred. No. 1.3e-11;		
Matches 64;	Conservative 27;	Mismatches 52;	Indels 18;	Gaps 5

```

Oy      36  PSTIVHCAVEIE-----ORGLTEGVIRISCSDRTVELEKEFLRYVTYDLSKVO---413
Db      1  PLIVEKCEVIEIETLYLAERGLQEGEIGYRVSQSGSRVKELEKRAFDKQDAPDLSLSEKWE 60

Oy      414 -DTHACSLCKDELRLKEPILLTFPLNRAFMFAA--ETTDGNSIAAMYQAV-GEIIOAN 466
Db      61 FDVHHVAVAGLLKLYLRLEPRLPIYDLYEEFFIAAKGEQLEDPERLRALKELSSKLPRAH 1200

Oy      470 RDTLAFLLMTHLORVAO-----SPHTRMDVANLAKVGPITY 505
Db      121 YNPLRLYLTHLNKNVAEITYENSAVKKMKNRNLAIAYFGPTLL 161

```

```

RESULT 13
US-09-764-868-898
: Sequence 898, Application US/09764868
: Patent No. US20020168711A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT232
: CURRENT APPLICATION NUMBER: US/09/764,868
: PRIORITY FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1510
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 898
: LENGTH: 291
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-764-868-898

```

Query Match	7.6%;	Score 247.5;	DB 9;	Length 291;
Best Local Similarity	30.0%;	Pred. No. 5.3e-11;		
Matches 89;	Conservative 39;	Mismatches 106;	Indels 63;	Gaps 13

```

QY 315 -----KRCDCR--VVSHEPCRD--RCPLPCIPILI-----GTPV 344
Db 58 SAAOTHOILRLRGAPACRECEAFMWSGTECECEFLTCHAKRILETLILLCGHRRILPARTPL 117
QY 345 KIGEGML--ADPVSOTSPMIPISIVHCNEIEOGRGLFHEGIVRISGCDRVKELKCKFL 401
Db 118 -FQVDLQLPRDPPEE---VFVVTKTAELHNHALVOGIVRYSGRSRAVENLCOAFE 172
QY 402 RKTIVBLKSVDDIHAICSLKDLFLRNKKEPLFLRLNARMEAAETIDED-----452
Db 173 NGRALVELSG-NSPHVSSVTLKRFQELTEPEYIPHLIDAFISLAKTLHADPGDGPNS 231
QY 453 ---NSTAAYQAVGELPQANRDTLAFIMIHQRV-AQSPHKKMYANLAKYFGTIV 505
Db 232 PSEPVRISLTKLLVQLPDSNRYNTLLHLVLAHFRVAARMENKMSANNGLIVGGTLL 288

```

```

RESULT 14
US-10-132-585-5
: Sequence 5, Application US/10132585
: Publication No. US20030055234A1
: GENERAL INFORMATION:
: APPLICANT: Kapeller, Libermann, Rosanna
: TITLE OF INVENTION: 26030, A HUMAN RHO-GAP FAMILY MEMBER AND
: TITLE OF INVENTION: US$ THEREFOR
: FILE REFERENCE: MP101-101PIRM
: CURRENT APPLICATION NUMBER: US/10/132,585
: CURRENT FILING DATE: 2002-04-25
: PRIOR APPLICATION NUMBER: 60/286,581
: PRIOR FILING DATE: 2001-04-25
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 152
: TYPE: PRT
: ORGANISM: unknown
: FEATURE:
: OTHER INFORMATION: Probom consensus rhogAP domain
US-10-132-585-5

```

Query Match	6.98;	Score 223.5;	DB 9;	length 152;
Best Local Similarity	39.38;	Pred. No. 1.4e-09;		
Matches 55;	Conservative 25;	Mismatches 39;	Indels 21;	Gaps 6

[illegible]

RESULT 15  
 US-10-080-960-34  
 ; Sequence 34, Application US/10080960  
 ; Publication No. US20020197695A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals, Inc.  
 ; APPLICANT: Glucksmann, Maria  
 ; APPLICANT: Meyers, Rachel  
 ; TITLE OF INVENTION: 80090, 52874, 52880, 63497, AND 33425  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 38155-20044, 00  
 ; CURRENT APPLICATION NUMBER: US/10/080,960  
 ; CURRENT FILING DATE: 2001-10-19  
 ; PRIOR APPLICATION NUMBER: US 60/242,040  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: US 60/242,038  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: US 60/241,992  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: US 60/242,637  
 ; PRIOR FILING DATE: 2000-10-23  
 ; NUMBER OF SEQ ID NOS: 37  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 34  
 ; LENGTH: 103  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Consensus amino acid sequence  
 ; US-10-080-960-34

Query Match	6.7%;	Score 218;	DB 9;	Length 103;
Best Local Similarity	45.1%;	Pred. No. 2.1e-09;		
Matches . 46;	Conservative 21;	Mismatches 29;	Indels 6;	Gaps 3;

409 LSKVDDIHAICSLKDFLRNKEPDLTFRNRAFMEA--AEITDEDSIAA---MYQAVG 463

Db 2 MEEYEDVHTVAGLLKQYFRELPEPLTTYELEYEEFIEAKAQVSDDEDERMEALEMUKELIK 61

QY 464 ELPOANDTLAFLMIHLQRYAQ-SPHTKMDVANLAKVEGPTI 504

Db 62 LPEANRETLLRYLLKHLSPAQHSEENKMAQNLAVEGPTL 103

Search completed: July 3, 2003, 10:05:54  
job time : 57 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:50:57 ; Search time 27 Seconds

(without alignments)  
688.714 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMMLNVRNLEFQVLRVRE.....SKSATNLGROGNFASPMUK 632

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/prodata/1/laa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/laa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/laa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/laa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/laa/6CTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/laa/6ackfilssl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	301	9.3	1261	3	US-09-080-855-2
2	270.5	8.3	2548	4	US-09-172-422-1
3	257.5	7.9	165	5	PCT-US93-03076-10
4	230	7.1	140	5	PCT-US93-03076-9
5	210	6.5	1513	5	PCT-US93-03076-2
6	191	5.9	2071	4	US-09-415-522-6
7	183.5	5.7	162	5	PCT-US93-03076-8
8	161	5.0	1013	4	US-09-415-522-8
9	128.5	4.0	878	4	US-09-735-934A-2
10	127.5	3.9	1354	3	US-08-685-871-2
11	126.5	3.9	724	1	US-07-906-349A-5
12	126.5	3.9	724	1	US-08-167-035-2
13	126.5	3.9	724	1	US-08-208-887A-2
14	126.5	3.9	724	2	US-08-539-005-2
15	126.5	3.9	724	2	US-09-280-598-5
16	125.5	3.9	722	3	US-08-390-874C-12
17	125.5	3.9	722	3	US-09-265-772-12
18	118	3.6	720	3	US-08-899-437-6
19	118	3.6	720	4	US-09-126-121-6
20	116.5	3.6	942	4	US-08-560-005-6
21	116.5	3.6	942	4	US-09-418-540-6
22	116	3.6	1146	4	US-08-914-999-6
23	115.5	3.6	666	4	US-09-134-001C-4739
24	114.5	3.5	1420	2	US-08-540-804-14
25	114.5	3.5	1420	2	US-08-218-265-14
26	114.5	3.5	1420	3	US-08-521-872-14
27	114.5	3.5	1420	4	US-08-590-399-14

28	114	3.5	894	4	US-09-735-934A-4	Sequence 4, App11
29	113.5	3.5	713	3	US-08-899-437-2	Sequence 2, App11
30	113.5	3.5	713	4	US-09-126-121-2	Sequence 2, App11
31	112.5	3.5	942	2	US-08-884-681-4	Sequence 4, App11
32	112.5	3.5	942	2	US-09-258-643-4	Sequence 4, App11
33	111	3.4	1003	1	US-08-571-758-4	Sequence 4, App11
34	111	3.4	1003	1	US-08-909-984A-4	Sequence 4, App11
35	111	3.4	1003	1	US-08-909-983-4	Sequence 4, App11
36	110	3.4	976	4	US-09-104-324B-4	Sequence 4, App11
37	107.5	3.3	1312	4	US-09-345-882-29	Sequence 29, App1
38	107.5	3.3	1471	3	US-08-755-887-188	Sequence 188, App
39	106.5	3.3	1618	1	US-07-853-913-4	Sequence 4, App11
40	105.5	3.3	588	4	US-08-235-836C-122	Sequence 122, App
41	105.5	3.3	700	1	US-07-720-589-2	Sequence 2, App11
42	105.5	3.3	700	2	US-08-785-190-2	Sequence 2, App11
43	105.5	3.3	700	4	US-08-235-836C-66	Sequence 66, App1
44	105.5	3.3	700	4	US-08-235-836C-74	Sequence 74, App1
45	105.5	3.3	700	5	PCT-US92-05539-2	Sequence 2, App11

## ALIGNMENTS

```

RESULT 1
US-09-080-855-2
; Sequence 2, Application US/09080855A
; Patent No. 6083721
; GENERAL INFORMATION:
; APPLICANT: Saras, Jan
; APPLICANT: Franz, Petra
; APPLICANT: Asperstrm, Pontus
; APPLICANT: Helman, Ulf
; APPLICANT: Gomez, Leonel Jorge
; APPLICANT: Heldin, Carl-Henrik
; TITLE OF INVENTION: PARC, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PP1A
; FILE REFERENCE: L0461/7030
; CURRENT APPLICATION NUMBER: US/09/080,855A
; EARLIER FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: 08/805,583
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-080-855-2

Query Match          9.3%; Score 301; DB 3; Length 1261;
Best Local Similarity 22.4%; Pred. No. 4e-19;
Matches 153; Conservative 120; Mismatches 249; Indels 160; Gaps 33;

OY      9  RLLEFQVLRVREISGNEVOFLOLAKDEDFRKKWQORDHE-IGRYKDLMLKAETTERSA 67
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      372 RLLEFQVLRVREISGNEVOFLOLAKDEDFRKKWQORDHE-IGRYKDLMLKAETTERSA 67
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY      68 LDYKTLKARNQVDVEIKRQRAADCEKLEKIQILREMLMCDTSSISLSEBOK---SA 124
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      428 VYVNLFFHMOHLAASLADRIQSLCGSAKYDPQOEVSEFVKATNS-----TEBEKVDGV 482
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY      125 LAFNLGQSSSSNAGKRLSTI---DESGTSLDISFDKTDSELDWSSLVYTFKLRKE 181
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      483 NKLHNSQPSGSGPANSLEDVYALPDSSKKITEEDRCSNSADIT---GPFISWTFGMS 539
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY      182 KRSTSRQFVDPGPPGPKRTKRSIGSAVDGNSIYAKTIVYVNDGPIEAVSTLETVPY 241
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      540 DSEST-----GSSSESRLDS-----EST-----SPGD----- 562
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY      242 WTRSRKKTGTLPQWNSDTLNSKOLEPRTET--DSVGTQSS---NGMRLHDFVSKTVLK 296
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      563 FHRKFLPRTSSSGTMSADLDEREPPSPETGNSIGTFRKTLMSKALVTHRF--RKLRS 620
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY      297 PESCVPCGRIKRIKFKLSLKCRCRCRVVSHPECRDRCLPC-----IP---TLIGPVIKIGSG 349
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db 621 PTKRDEGLVFE--QCEVECECLVCHRCLEMLVITICGHOXPKIHLFG----- 670  
QY 350 MLADP--VSQTSR--MISIVVHCNVEIEORGLTETGLYRISGCDRTVKELEKFLRYKTV 406  
Db 671 --AEFTLVAKKEPGICIFILKICASELEENRACLOGIYRVC-----NKIKTE 716  
QY 407 PLASKVD-----DI-----HAISLKLDFLNKEPILFRLNRAFMKAE-----ITDE 451  
Db 717 KLCLALENGMLVLDISESSHDICDVLKYLRLQPEPILFRLKEPILDAKEIQHNEE 776  
QY 452 DNS-----IAAMYQAVGELPQANRDLALMLHLQRYA--QSPHTK 490  
Db 777 QETKNSLEDKKMNMCIEINRLKSKDLRLQPLASNFSNLHLVHLKRVVDHAEENK 836  
QY 491 MDVANLAKVGPRTIYAAVNPDP--PYTMSQ--DIRKOPKYVERLSLPLEYMSQFM--- 543  
Db 837 MNSKNLGVIFGPISLR---FRPQIAPITISSIATYMSQARVLEPLT---YSOKIFDG 888  
QY 544 --MVEQENIDPLHVIENS--NAFSTPQTPDIKVSLLGPVTPPEHOLKTPSSSLSQVR 599  
Db 889 SLQPDVWCSIGVYDQCFPKPLSPERDIERSMKSLFSSKEDIHTSESEKIFERAT 948  
QY 600 STLTKNTPRGSKSKSATNIGR 621  
Db 949 S-----FEESERKONALGK 962

## RESULT 2

US-09-172-422-1  
; Sequence 1, Application US/09172422A  
; Patent No. 6300485  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Arwen E.  
; APPLICANT: Chiu, Choi Yung  
; APPLICANT: Duhl, David  
; APPLICANT: Gorman, Susan W.  
; APPLICANT: Leng, Song  
; APPLICANT: Sheffield, Val  
; APPLICANT: Welch, Juliet  
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED  
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,  
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF  
; FILE REFERENCE: 200130.442  
; CURRENT APPLICATION NUMBER: US/09/172.422A  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2548  
; TYPE: PRT  
; ORGANISM: Homo sapien  
; US-09-172-422-1

Query Match 8.3%; Score 270.5; DB 4; Length 2548;  
Best Local Similarity 24.4%; Pred. No. 1e-15;  
Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

QY 143 LSTIDSGSLIDSEPDKTESLMDSSLVKTFKLKRRKRSRSRPFVGPPEPVKTR 202  
Db 1864 LKSDDE--FLTKRVN--DLDNEDSKKDLVDVVEFKALKERON-----IPSFY 1908  
QY 203 SIGSAVDQGN-----ESIVAKTYTVPNMG--GPIEA-VSTIETVPWTRSR 247  
Db 1909 SSALAMDGKSIRKDLALFEOLETMTMLEQDSDGSESPVRKRWVTFKVLDEYNNF 1968  
QY 248 KTGTLQFWSNSTL-----NSROLEPTEFDSVGTPOSGMRMLHDFSVSKTVIKPESCV 302  
Db 1969 KT-----SDCATKVPKTERKKRRKRETDLV--EEHNG---HIFKATQYSIPTVCEY 2015  
QY 303 CGKRIKGTSLKSCDRCRVVSHPECRCRCLPCLPTLIGTPVKG-----GMLADFY 355  
Db 2016 CSSLIWIMDRASVCKLCKYACHKCC-----CIKTKAKCSKYPDELSSROFGVELSRL 2068

QY 356 SQSPMPSIVVHCNVEIEORGLTETGLYRISGCDRTVKELEKFLRYKTVPLLSKYD- 414  
Db 2069 TSEDRTPVPLVEKININIEHGLYTEISYKSSGTNKKIKELRO---GLDIDASVNLDDY 2125  
QY 415 --HAISLKLDFLNKEPILFRLNRAFMKAEITDEDSIAAMYQAVGELPQANRDL 473  
Db 2126 NIHVIAVSFQWMLDLPNPLMTFELYEEFLRAMLOERKREKTIQVSVIQLSRTHLNTL 2185  
QY 474 AFLMIHLQRYA--QSPHTKMNVANLAKVGPRTIYAAVNPDPYTMQDIDKROKRYERL 532  
Db 2186 ERLFLHLYRIALQEDTRKMSANALAIYFACIL--RCDDTIDPLQSDVODISKTTCV- 2241  
QY 533 SLPLEYMSQFMVEQEN-----IDPLHVIENSNAFSTPQTPDIKVSLL----- 575  
Db 2242 -----LIVQNMKRYKARLKDLSLFEAEK-----AKTRLSLIRSMGKRI 2284  
QY 576 -----GPVTPPEHOLKTPSSSLSQVRVSTLTKNTPRGSKSKSATNIGRQ 622  
Db 2285 RRGNYRGP--SSP--VVVRLPVSQVSE---ETLT-----SEAMETDITEQ 2324

## RESULT 3

PCT-US93-03076-10  
; Sequence 10, Application PC/TUS9303076  
; GENERAL INFORMATION:  
; APPLICANT: Whitehead Institute for Biomedical Research  
; TITLE OF INVENTION: GAP-Associated Protein p190 and  
; TITLE OF INVENTION: Transduction  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: 2 Millitia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/03076  
; FILING DATE: 19930331  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: WHI92-03A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 165 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US93-03076-10

Query Match 7.9%; Score 257.5; DB 5; Length 165;  
Best Local Similarity 37.9%; Pred. No. 1.8e-16;  
Matches 64; Conservative 33; Mismatches 63; Indels 9; Gaps 4;

QY 369 CVNEIEORGLTETGLYRISGCDRTVKELEKFLR--VKTVPLLSKYVDIHAISLKLDFL 426  
Db 1 CIREIESRGINSBGLYVSGFSDLIBDVKNAFPDGEKADISVNMVEDINITIGALKIYF 60  
QY 427 RNKKEPLTFLRLNRAFMKAEITDEDSIAAMYQAVGELPQANRDLALMLHLQRYAOS 486  
Db 61 RDLPIPLITYDAVPKFLIESAKIMDPDEOLETLHEALAKLLPAPACETLRYLMAHLKRV--T 118









Db 1085 -----SDIEOLRAKLIDLSSTSVASEPSSADEFDGNLP--ESRIEGMLSVNPRGNIK 1134  
QY 183 RRSRSROFVGGPPGPKTKRSI-----GSANDQNESTIVA---TTVTVPDGGPIEAYS 234  
Db 1135 RYCKKKQYV-----VSSSKILFYNDQDKEQSNPSMVLIDILFHVPRPTQGDVYRAE 1188  
QY 235 TIEVTPWTRSRKRTGTLOPWNDSSTINSROLEPRTETDSVGTPOSGMRLHDHVFSEKTV 294  
Db 1189 T-EEIP-----KIFDILVANBEGCKRDVEMEPVQAETKNPNHNG----HEFIFPLY 1236  
QY 295 IKPESCVPCGK---RIFGKILSKCQRCRVVSHPECRCDCPLPCIPPLIGTPVKIGEGML 351  
Db 1237 HFPAACDACAQKPLMHVFKPPPALECRCHYCHRDHLDKEDLIPCCKVSYDTYSARDML 1296  
QY 352 ADFVSQTSPIHPSIVHVCNVEIEQRLTETGLYRIS 387  
Db 1297 LMACSODEO--KKWVTHLVKKIPKN--PPSGFVNAS 1328

## RESULT 11

US-07-906-349A-5  
; Sequence 5, Application US/07906349A  
; Patent No. 5434064

## GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnick, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
TITLE OF INVENTION: A NOVEL EXPRESSION CLONING METHOD FOR  
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN  
NUMBER OF SEQUENCES: 16  
TITLE OF INVENTION: TARGET PROTEINS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Neimark  
STREET: 419 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,349A  
FILING DATE: 30-JUN-1992  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/643,237  
FILING DATE: 18-JAN-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528

## INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 724 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-906-349A-5

Query Match 3.9%; Score 126.5; DB 1; Length 724;  
Best Local Similarity 22.2%; Pred. No. 0.0054;

Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

QY 325 PECDRCPLPCIPPLITLGTVPKIGEGMLA--DFVSQTSPP--MISIVVHCNVEIEQRLTE 380  
Db 87 PKRPPRPLPVAVGSSKTEADVEQALTLPLDLAEQAFAPDIPAPLILKIVAEIEKKGLGC 146  
QY 381 TGLYRISGCRVYKELKEKFLRYKTVPLLSKVDIDIAICSLDDELNLKEPLTLFRNLN 440

Db 147 STYRTQS--SSNLAELRQ--LLDCPTPSVDELMIDVHVLADAFKRYLLDLDPNPIPAVYS 204  
QY 441 AFME-AAETIDEDNSIAAMYQAV--GELPOANDRTLAFLMIHQVRAQ--SPHTKMDVANTL 496  
Db 205 EMISLAPEVOSSEYIDLLKLRSPSPIDHQLTLQYLKHFPKLSQTSKMLLNARVL 264  
QY 497 AKVGPPIVAHAVPNDPVTMSODIKRQPKVVERLSLPLEYVSQ 541  
Db 265 SEIFSPMLFRSAASSD-----NTENLIVIELISTE---WNE 300

## RESULT 12

US-08-167-035-2  
; Sequence 2, Application US/08167035  
; Patent No. 5618691

## GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnick, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR  
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE  
KINASES AND NOVEL TARGET PROTEINS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: New York  
ZIP: 10036-2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,035  
FILING DATE: 16-DEC-1993  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 724 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-167-035-2

Query Match 3.9%; Score 126.5; DB 1; Length 724;  
Best Local Similarity 22.2%; Pred. No. 0.0054;

Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

QY 325 PECDRCPLPCIPPLITLGTVPKIGEGMLA--DFVSQTSPP--MISIVVHCNVEIEQRLTE 380  
Db 87 PKRPPRPLPVAVGSSKTEADVEQALTLPLDLAEQAFAPDIPAPLILKIVAEIEKKGLGC 146  
QY 381 TGLYRISGCRVYKELKEKFLRYKTVPLLSKVDIDIAICSLDDELNLKEPLTLFRNLN 440  
Db 147 STYRTQS--SSNLAELRQ--LLDCPTPSVDELMIDVHVLADAFKRYLLDLDPNPIPAVYS 204  
QY 441 AFME-AAETIDEDNSIAAMYQAV--GELPOANDRTLAFLMIHQVRAQ--SPHTKMDVANTL 496  
Db 205 EMISLAPEVOSSEYIDLLKLRSPSPIDHQLTLQYLKHFPKLSQTSKMLLNARVL 264  
QY 497 AKVGPPIVAHAVPNDPVTMSODIKRQPKVVERLSLPLEYVSQ 541



